GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 1, 1999, 15:34:36 ; Search time 1619.02 Seconds
(without alignments)
45.180 Million cell updates/sec

Title: US-09-026-400-5
Perfect score: 23
Sequence: 1 gcngtngartggaayttygcnmg 23

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database :
GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_st:*
14: gb_sts:*
15: gb_sy:*
16: gb_un:*
17: gb_vi:*
18: em_fun:*
19: em_htg:*
20: em_hum1:*
21: em_hum2:*
22: em_in:*
23: em_om:*
24: em_or:*
25: em_ov:*
26: em_pat:*
27: em_ph:*
28: em_pl:*
29: em_ro:*
30: em_sts:*
31: em_sy:*
32: em_un:*
33: em_vi:*
34: gb_htg1:*
35: gb_htg2:*
36: gb_in1:*
37: gb_in2:*
38: em_ba1:*
39: em_ba2:*
40: em_hum3:*
41: em_hum4:*
42: gb_pr4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	15.8	68.7	205	9	HS164H1R	Z54791 H.sapiens C
2	15.8	68.7	1721	11	HSU50733	U50733 Human dynam
c 3	15.8	68.7	25321	36	CEF18E2	Z75537 Caenorhabdi
4	15.6	67.8	2061	3	BTA6304	AJ006304 Bos tauru
c 5	15.6	67.8	2716	7	CSLYS4	X93502 S.cerevisia
c 6	15.6	67.8	30507	7	SC8419	Z49701 S.cerevisia
c 7	15.6	67.8	2710	8	SCU46154	U46154 Sacchoromyc
8	15.6	67.8	4564	10	HSAJ6276	AJ006276 Homo sapi
9	15.6	67.8	353	10	HSTRPC3GN	X89068 H.sapiens m
10	15.6	67.8	3448	10	HSY13758	Y13758 Homo sapien
11	15.6	67.8	151630	11	AC005154	AC005154 Homo sapi
12	15.6	67.8	147216	11	AF093117	AF093117 Homo sapi
13	15.6	67.8	3417	11	HSU47050	U47050 Human putat
14	15.6	67.8	2842	12	AF057748	AF057748 Mus muscu
15	15.6	67.8	3691	12	AF111107	AF111107 Mus muscu
16	15.6	67.8	3608	12	AF111108	AF111108 Mus muscu
17	15.6	67.8	3340	12	AF136401	AF136401 Rattus no
18	15.6	67.8	3261	12	MMU49069	U49069 Mus musculu
19	15.6	67.8	124156	35	AC006969	AC006969 Homo sapi
20	15.6	67.8	448	36	CETRPHOMO	X90698 C.erythroce
21	15.6	67.8	3867	36	CVTRP	Z80230 Calliphora
c 22	15.6	67.8	86603	37	AC005120	AC005120 Drosophil
c 23	15.6	67.8	129779	37	AC006214	AC006214 Drosophil
24	15.6	67.8	2796	42	AF080394	AF080394 Homo sapi
25	15.4	67.0	38052	10	HSAC000363	AC000363 Human cos
c 26	15.4	67.0	76416	11	AC005370	AC005370 Homo sapi
c 27	15.4	67.0	110000	34	CEY37H9_1	Continuation (2 of
c 28	15.4	67.0	87805	34	CEY37H9_3	Continuation (4 of
c 29	15.4	67.0	110000	34	CEY71A12_0	AL021390 Caenorhab
c 30	15.4	67.0	33455	36	CEW04A8	Z82069 Caenorhabdi
c 31	15.2	66.1	2333	1	KPNDEAD	L08387 Klebsiella
32	15.2	66.1	37114	1	MSGB38COS	L01095 M. leprae g
c 33	15.2	66.1	715	4	GGU09261	U09261 Guira guira
c 34	15.2	66.1	1045	4	SCU37304	U37304 Synthlibora
c 35	15.2	66.1	1143	4	U89204	U89204 Pharomachru
36	15.2	66.1	71618	7	AB012240	AB012240 Arabidops
37	15.2	66.1	114367	7	ATF10M10	AL035521 Arabidops
c 38	15.2	66.1	94091	7	ATF28A23	AL021961 Arabidops
c 39	15.2	66.1	4748	7	PSU51562	U51562 Prunus sero
40	15.2	66.1	6217	9	AB014606	AB014606 Homo sapi
41	15.2	66.1	84170	9	HS130N4	Z75887 Human DNA s
42	15.2	66.1	44888	9	HS130N4B	Z75888 Human DNA s
43	15.2	66.1	288	9	HS44A5F	Z58565 H.sapiens C
c 44	15.2	66.1	276	9	HS44A5R	Z58566 H.sapiens C
45	15.2	66.1	37738	36	CELF32B5	AF003148 Caenorhab

ALIGNMENTS

RESULT 1
HS164H1R
LOCUS HS164H1R 205 bp DNA PRI 17-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 164h1, reverse read cpg164h1.rt1a.
ACCESSION Z54791
NID g1020832
VERSION Z54791.1 GI:1020832
KEYWORDS CpG island; genomic MseI fragment.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 205)
AUTHORS Dodsworth,S.J., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
REFERENCE 2 (bases 1 to 205)
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
COMMENT Vector: pGEM-5Zf(-)

```
          Clones are available from the UK MRC Human Genome Mapping Project
          Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
          http://www.hgmp.mrc.ac.uk/ for details
          or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .205
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /sex="male"
                     /dev_stage="adult"
                     /tissue_type="blood"
                     /clone_lib="CGI-1"
                     /clone="164h1"
BASE COUNT       49 a     45 c     63 g     47 t      1 others
ORIGIN
```

```
  Query Match             68.7%;  Score 15.8;  DB 9;  Length 205;
  Best Local Similarity   63.6%;  Pred. No. 3.4e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 gcngtngartggaayttygcnm 22
           || || ||:|| ||:||:|| :
Db      52 GCAGTGGAGTGAAACTTTGCCC 73
```

```
RESULT   2
HSU50733
LOCUS       HSU50733     1721 bp    mRNA            PRI       04-APR-1996
DEFINITION  Human dynamitin mRNA, complete cds.
ACCESSION   U50733
NID         g1255187
VERSION     U50733.1  GI:1255187
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1721)
  AUTHORS   Echeverri,C.J., Paschal,B.M., Vaughan,K.T. and Vallee,R.B.
  TITLE     Molecular characterization of the 50-kD subunit of dynactin reveals
            function for the complex in chromosome alignment and spindle
            organization during mitosis
  JOURNAL   J. Cell Biol. 132 (4), 617-633 (1996)
  MEDLINE   96178072
REFERENCE   2  (bases 1 to 1721)
  AUTHORS   Echeverri,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAR-1996) Christophe J. Echeverri, Cell Biology,
            Worcester Foundation for Biomedical Research, 222 Maple Avenue,
            Shrewsbury, MA 01545, USA
FEATURES             Location/Qualifiers
     source          1. .1721
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             79. .1299
                     /note="similar to GenBank EST Accession Number T08494; p50
                     subunit of dynactin complex"
                     /codon_start=1
                     /product="dynamitin"
                     /protein_id="AAC50423.1"
                     /db_xref="PID:g1255188"
                     /db_xref="GI:1255188"
                     /translation="MADPKYADLPGIARNEPDVYETSDLPEDDQAEFDAFAQELEELT
                     STSVEHIIVNPNAAYDKFKDKRVGTKGLDFSDRIGKTKRTGYESGEYEMLGEGLGVKE
                     TPQQKYQRLLHEVQELTTEVEKIKTTVKESATEEKLTPVLLAKQLAALKQQLVASHLE
                     KLLGPDAAINLTDPDGALAKRLLLQLEATKNSKGGSGGKTTGTPPDSSLVTYELHSRP
                     EQDKFSQAAKVAELEKRLTELETAVRCDQDAQNPLSAGLQGACLMETVELLQAKVSAL
                     DLAVLDQVEARLQSVLGKVNEIAKHKASVEDADTQSKVHQLYETIQRWSPIASTLPEL
                     VQRLVTIKQLHEQAMQFGQLLTHLDTTQQMIANSLKDNTTLLTQVQTTMRENLATVEG
                     NFASIDERMKKLGK"
BASE COUNT       473 a    445 c    453 g    350 t
ORIGIN
```

```
  Query Match             68.7%;  Score 15.8;  DB 11;  Length 1721;
  Best Local Similarity   63.6%;  Pred. No. 2.8e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 cngtngartggaayttygcnmg 23
           | || ||: ||||:||:|| :|
Db    1244 CAGTTGAGGGGAACTTTGCCAG 1265
```

```
RESULT   3
CEF18E2/c
LOCUS       CEF18E2     25321 bp    DNA             INV       23-NOV-1998
DEFINITION  Caenorhabditis elegans cosmid F18E2, complete sequence.
ACCESSION   Z75537
NID         g1418504
VERSION     Z75537.1  GI:1418504
KEYWORDS    HTG; ABC transporter; Acid phosphatase like.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
            Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 25321)
  AUTHORS   Lightning,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-1996) Louis, MO 63110, USA. E-mail:
            jes@sanger.ac.uk or rw@nematode.wustl.edu
REFERENCE   2  (bases 1 to 25321)
  AUTHORS   Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M.,
            Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,
            Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,
            Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,
            Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N.,
            Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B.,
            O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A.,
            Saunders,D., Shownkeen,R., Smaldon,N., Smith,A., Sonnhammer,E.,
            Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M.,
            Vaughan,K., Waterston,R., Watson,A., Weinstock,L.,
            Wilkinson-Sproat,J. and Wohldman,P.
  TITLE     2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
            elegans
  JOURNAL   Nature 368 (6466), 32-38 (1994)
  MEDLINE   94150718
COMMENT     Coding sequences below are predicted from computer analysis, using
            predictions from Genefinder (P. Green, U. Washington), and other
            available information.
            For a graphical representation of this sequence and its analysis
            see:-
            http://webace.sanger.ac.uk/cgi-
            bin/display?db=wormace&class=Sequence &object=F18E2
            Current sequence finishing criteria for the C. elegans genome
            sequencing consortium are that all bases are either sequenced
            unambiguously on both strands, or on a single strand with both a
            dye primer and dye terminator reaction, from distinct subclones.
            Exceptions are indicated by an explicit note.
            IMPORTANT:  This sequence is NOT necessarily the entire insert of
            the specified clone.  It may be shorter because we only sequence
            overlapping sections once, or longer because we arrange for a small
            overlap between neighbouring submissions.
            IMPORTANT: This sequence is not the entire insert of clone F18E2.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we arrange for a small overlap between
            neighbouring submissions.
            The true left end of clone F18E2 is at 1 in this sequence. The true
            right end of clone F18E2 is at 14886 in
            sequence Z75554.
            The true left end of clone ZC455 is at 25221 in this sequence. The
            true right end of clone C15C8 is at 24380 in this sequence. The
            start of this sequence (1. .104) overlaps with the end of sequence
            Z75527.
            The end of this sequence (25221. .25321) overlaps with the start of
            sequence Z75554.
```

```
FEATURES             Location/Qualifiers
     source          1. .25321
                     /organism="Caenorhabditis elegans"
                     /db_xref="taxon:6239"
                     /chromosome="V"
                     /clone="F18E2"
     gene            complement(4069. .6129)
                     /gene="F18E2.1"
     CDS             complement(join(4069. .4165,4214. .4473,4519. .4653,
                     4702. .4812,4962. .5192,5250. .5497,5794. .5937,5988. .6129))
                     /gene="F18E2.1"
                     /note="Similarity to Aspergillus acid phosphatase
                     (TR:G755244)"
                     /codon_start=1
                     /protein_id="CAA99834.1"
                     /db_xref="PID:e1345818"
                     /db_xref="PID:g3876094"
                     /db_xref="GI:3876094"
                     /db_xref="SPTREMBL:Q19553"
                     /translation="MSIHITLTVISLLVTKCISEGTTPDQVHISFTGDMTEMAVVWNT
                     FSEVSQDVTYGKTGSGATSTAKGSSEAWVFGGITRYRHKAIMTGLEYSTEYDYTIASR
                     KFSFKTLSNDPQSYKVCVFGDLGYWHGNSTESIIKHGLAGDFDFIVHLGDIAYDLHTN
                     NGQVGDSYLNVFEPLISKVPYMVIAGNHEDDYQNFTNYQKRFSVPDNGHNDNQFYSFD
                     LGPVHWVGVSTETYGYYYEYGMDPVMTQYDWLKRDLTTANSNRAAHPWIFTFQHRPFY
                     CSNVNSAECQSFENRLVRTGWLDMPGLEPLFLQTSVDFGFWGHEHSYERFYPVADRAY
                     WNDPNAYINPKAPVYLISGSAGCHTPDALFTDKPWPWSAARNNDYGWSIVTVANRTHI
                     RVEQISIDKNEQTVDDFWVIKDEGHMHSGEMRRANPGAKFPLQKCHFKDVACKRSLKE
                     INEEL"
     gene            complement(7557. .9729)
                     /gene="F18E2.2"
     CDS             complement(join(7557. .8057,8105. .8669,8825. .9283,
                     9331. .9642,9698. .9729))
                     /gene="F18E2.2"
                     /note="Similarity to Yeast ATP-dependent transporter
                     (SW:YEM6_YEAST); cDNA EST EMBL:C09155 comes from this
                     gene; cDNA EST yk259a9.3 comes from this gene; cDNA EST
                     yk259a9.5 comes from this gene; cDNA EST yk453c2.3 comes
                     from this gene; cDNA EST yk453c2.5 comes from this gene;
                     cDNA EST yk454c10.5 comes from this gene; cDNA EST
                     yk456g7.3 comes from this gene; cDNA EST yk456g7.5 comes
                     from this gene; cDNA EST yk458h5.3 comes from this gene;
                     cDNA EST yk458h5.5 comes from this gene; cDNA EST
                     yk466h9.3 comes from this gene; cDNA EST yk504g10.3 comes
                     from this gene"
                     /codon_start=1
                     /protein_id="CAA99835.1"
                     /db_xref="PID:e1345819"
                     /db_xref="PID:g3876095"
                     /db_xref="GI:3876095"
                     /db_xref="SPTREMBL:Q19554"
                     /translation="MSDAGKKISRKEAKKAKAKENYDKTILSMGGVVENLPEVEGDQE
                     EHGGIGSGAELGAHFTVSQLSKTGTQLAQMENSMDIKIENFDISAQGKLLFDKASLTI
                     VYGRRYGLVGPNGMGKTTLLKHIGARKLAIPSHIDLLYCEQEIQVDSTSAIDTVVKSD
                     KKRLALLEEEAKLMSEIEEGKTEAAERMKEVADELRDIGADSAEPRARRILAGLGFSK
                     EMQEKPCTDFSGGWRMRISLARALFLEPTLLMLDEPTNHLDLNAVIWLDNYLQTWKKT
                     LLIVSHDQGFLDSVCTDIIHLDNQKLHTYRGNYTLFKKQYAQDMQVHEKNFDQQQKQL
                     KAMKKEGKSAKQAEEQVKQQMANKAKKGGKKNAGKVNDDDDAGAPELLQRRKEYSVKF
                     QFPETTKLNPPVLGLHDVNFGYGKDVLFKKLNFGVDMDSRIAIVGPNGVGKSTLLKLL
                     IGKIDPNDGELRKHRTLRIGWFDQHANEALNGEQTPVEFLCTKFNIDYQEARKQLGTT
                     GLAAHAHTVKIKDLSGGQKSRVALCNLALGGPDIIILDEPTNNLDIESIDALAEAIRD
                     FNGGVVMVTHDERLVVRTDCNLWVVENQGIDEIDGDFEDYKKEVLDALGEALVSKK"
     gene            10314. .14040
                     /gene="F18E2.3"
     CDS             join(10314. .10411,10459. .10866,10920. .11998,12234. .12520,
                     12568. .13757,13812. .14040)
                     /gene="F18E2.3"
                     /note="cDNA EST EMBL:T00548 comes from this gene; cDNA EST
                     EMBL:D33077 comes from this gene; cDNA EST EMBL:D33596
                     comes from this gene; cDNA EST EMBL:D35819 comes from this
                     gene; cDNA EST EMBL:D36799 comes from this gene; cDNA EST
                     EMBL:D36456 comes from this gene; cDNA EST EMBL:D72367
                     comes from this gene; cDNA EST EMBL:D75215 comes from this
                     gene; cDNA EST EMBL:D64592 comes from this gene; cDNA EST
                     EMBL:D67647 comes from this gene; cDNA EST yk448f1.3 comes
                     from this gene; cDNA EST yk448f1.5 comes from this gene;
                     cDNA EST yk260g1.3 comes from this gene; cDNA EST
                     yk260g1.5 comes from this gene; cDNA EST yk254d3.5 comes
                     from this gene; cDNA EST yk427e9.5 comes from this gene"
                     /codon_start=1
                     /protein_id="CAA99836.1"
                     /db_xref="PID:e1345820"
                     /db_xref="PID:g3876096"
                     /db_xref="GI:3876096"
                     /db_xref="SPTREMBL:Q19555"
                     /translation="MSETPTDQSPQRMSTRNQARVNYTDMAAGNNSVEKEPVFRSPTA
                     STRGRKKRAANVDVADLSASFGNLNNFSTPPKRGRPRGGGLGSARGGRAPMVRRTTTE
                     ESAEVDDRELVAAVKSGKKITEAVDRWIGRYNEKFLVAIAEMHQFFFAICGCKGIVTP
                     QMSATLTYKDIICRMTEDFEEDSADYPLVHGGSLKKVRANLHNFIHTLIIRTKASMLF
                     DSNLMDGFVQLLTGMADSQVRAFRHTATFCAMKITSALVDVTIELTQSKEKTSKQIEA
                     EKAKLKNNSAGNEKYEALVAQRTQTEERAEEIRQIIGYLFRSVFVHRYRDVVPDIRCI
                     CIQELGHWMDVYPEHFVEDSYLKYIGWSMFDKVGDVRQRCIRALIPLFEKTLILDKLE
                     LFVNKFKDRLVSMLLDKDLETSIETVHLMRVLYTVFPTLLTIKDVVPIYELIYASNRP
                     LAVAAGMFLNTKVFRSAEKPGKAPTAANIPLVKDLTTFFIEGDLHQHATYLVDALFES
                     NPIVKDWATMGELLINDQYQLDSNFETKLIEILTCSVVQSATGEPPVGRHIVKKGAPS
                     AKESRDLVEDRQRLTETLIPLIPRLITKFSSDNEKIINLVNIPLHFQLETYLSARMQT
                     HLMELMEGLDSLIEKHLDEELLKAVAELYYHLTTNSSISALVEGHKMKLLDGVAAFIR
                     KSMQQFDDDQMGEEEEALFVSYIKRMAAFAGFMDLRHWDLWDILLKVVSNYTREDTQR
                     DVRERSMQMLFMQLCFDSMNIKKEGETPKADQVRKLKKRRDQLIRIVTETLNEEACGV
                     EQAYLVICDLMILFGSQLAEESKALEPLIWRPDAMVLGNLKIFLDVNVFDVSNLDDMD
                     QQEKIEVMHKMRQHVAQYAKLIIHGAMPVAEASHLIKRYQSHFQDFGDIFKNLLSKCR
                     EISFVETGVMICETLKTLYSQLDEDQGTDPLSESFNSIRDLAKRLGPAFGVDYAKNRF
                     AISSLHKKAIDFAFEEYDKENHQMPKNIFFLEIAIEFSGKLLAQDKMAVVRYLNKIYT
                     NRVGTSTVVWEPYRLYLGSLSDRNDDDNMSVRSGMTVTSNATMRSTASSTRGRGRGRG
                     RSRIADDF"
     gene            15418. .16931
                     /gene="F18E2.4"
     CDS             join(15418. .15456,15622. .15886,15933. .16028,16166. .16419,
                     16533. .16931)
                     /gene="F18E2.4"
                     /codon_start=1
                     /protein_id="CAA99837.1"
                     /db_xref="PID:e1345821"
                     /db_xref="PID:g3876097"
                     /db_xref="GI:3876097"
                     /db_xref="SPTREMBL:Q19556"
                     /translation="MYTVVSIHRCSLLTIFFNMNRIIQYGSVEGIPYYNCSAKTSEEW
                     FATGVQRPLFGWSILIFGIVIELLYIPTIYMMFRTKLIHLTCYKIIVCLGVTDMLATL
                     TCSIFSGWLFIKGAVYCNYPTFIYLAGCLGLSGWCMACGSTLLLVINRVFDVLCRPVS
                     EFLFDGKRVFFSIALIYTYGTLVSMFSPSIIFNSAIMAWIADPLTIDPEYKTDEITDW
                     YRNHVQSTNNWIFCSGTCTLYAIYCVLINKMQRGQKSKASRQVFIQSSIICSFNTFSA
                     MSYNALMFTVPAPWVIAIAELCWSIVHGCPAIIYLTMNKTIRTEFLKFLKLKKTNKVN
                     NATSFSNMNTSKSNFL"
     gene            join(complement(1233. .1262),complement(816. .909),
                     complement(578. .762),complement(46. .200),
                     complement(Z75527.1:20448. .20576),
                     complement(Z75527.1:20263. .20392),
                     complement(Z75527.1:20063. .20216),
                     complement(Z75527.1:19840. .19976),
                     complement(Z75527.1:19578. .19634))
                     /gene="gpa-13"
     CDS             join(complement(1233. .1262),complement(816. .909),
                     complement(578. .762),complement(46. .200),
                     complement(Z75527.1:20448. .20576),
                     complement(Z75527.1:20263. .20392),
                     complement(Z75527.1:20063. .20216),
                     complement(Z75527.1:19840. .19976),
                     complement(Z75527.1:19578. .19634))
                     /gene="gpa-13"
                     /note="Similarity to Drosophila Guanine nucleotide-binding
                     protein (SW:GB01_DROME)"

  Query Match          68.7%;  Score 15.8;  DB 36;  Length 25321;
  Best Local Similarity  63.6%;  Pred. No. 2.2e+02;
  Matches  14;  Conservative   4;  Mismatches   4;  Indels   0;  Gaps   0;

Qy       1 gcngtngartggaayttygcnm 22
```

```
          || || ||:| |||:||:|| :
Db   3680 GCTGTTGAATTGAATTTTGCAA 3659


RESULT   4
BTA6304
LOCUS       BTA6304      2061 bp    mRNA            MAM       11-JUN-1998
DEFINITION  Bos taurus mRNA for TRP2 protein.
ACCESSION   AJ006304
NID         g3152951
VERSION     AJ006304.1  GI:3152951
KEYWORDS    TRP-related; TRP2 gene; Trp2 protein.
SOURCE      Bos taurus.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Artiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 2061)
  AUTHORS   Wissenbach,U.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-MAY-1998) Wissenbach U., Universitat des Saarlandes,
            Pharmakologie und Toxikologie, D-66421 Homburg, Saarland, GERMANY
REFERENCE   2  (bases 1 to 2061)
  AUTHORS   Wissenbach,U., Schroth,G., Philipp,S. and Flockerzi,V.
  TITLE     Structure and mRNA expression of a bovine trp homologue related to
            mammalian trp2 transcripts
  JOURNAL   FEBS Lett. 429 (1), 61-66 (1998)
  MEDLINE   98319466
FEATURES             Location/Qualifiers
     source          1. .2061
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /tissue_type="testis"
     gene            613. .1911
                     /gene="trp2"
     CDS             613. .1911
                     /gene="trp2"
                     /codon_start=1
                     /product="tyrosine-related protein 2"
                     /protein_id="CAA06964.1"
                     /db_xref="PID:e1294008"
                     /db_xref="PID:g3152952"
                     /db_xref="GI:3152952"
                     /db_xref="SPTREMBL:O62826"
                     /translation="MVILSLYLAAFTLRLLLAGLAHKHCRDAPDGAACHYFTSAERSE
                     WRTEDPQFLAEVLFAVTSMLSFTRLASILPAHESLGTLQISMGRMIDDMIRFMFILMI
                     ILTAFLCGLNNIYVPYQETERLGNFNETFQFLFWTMFGMEEHSVVDMPQFLVPEFVGR
                     ALYGIFTIVMVIVLLNMLIAMITNSFQKIEDAADVEWKFARSKLYLSYFREGLTLPVP
                     FNILPSPKAIFYLLRRVFRFICCCHFCCKTKKPDYPPIPTFANPGAGAGPGEGERGSY
                     RLRVIKALVQRYIETAQREFEETRRKDLGNRLTELTKTVSRLQSEVAGVQRAVVEAGP
                     RRPPGGASVLSRYITRVRNSFQNLGPPIPETPELTVPATVGTQESSEIGLPDAGGAQA
                     PASGESGPSSPAHVLVHREQESEGAGDLPQEADLGAKEGT"
BASE COUNT      412 a    604 c    592 g    453 t
ORIGIN


  Query Match             67.8%;  Score 15.6;  DB 3;  Length 2061;
  Best Local Similarity   65.2%;  Pred. No. 3.5e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 gcngtngartggaayttygcnmg 23
          |  || ||:||||| ||:|| :|
Db   1192 GACGTGGAGTGGAAGTTTGCTCG 1214


RESULT   5
CSLYS4/c
LOCUS       CSLYS4       2716 bp    DNA             PLN       28-NOV-1995
DEFINITION  S.cerevisiae LYS4 gene.
ACCESSION   X93502
NID         g1089839
VERSION     X93502.1  GI:1089839
KEYWORDS    homoaconitase; Lys4 gene.
SOURCE      baker's yeast.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
            Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 2716)
  AUTHORS   Gamonet,F. and Lauquin,J.M.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2716)
  AUTHORS   Gamonet,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-NOV-1995) F. Gamonet, IBGC-CNRS, CNRS, 1, rue Camille
            Saint Saens, F- 33077 Bordeaux Cedex, FRANCE
FEATURES             Location/Qualifiers
     source          1. .2716
                     /organism="Saccharomyces cerevisiae"
                     /strain="GRF88"
                     /db_xref="taxon:4932"
                     /chromosome="4"
                     /clone_lib="CEN-ARS YCp50"
                     /clone="pV2"
                     /map="right arm"
                     /note="isogenic to strain S288-C"
     sig_peptide     277. .336
                     /gene="lys4"
                     /db_xref="SGD:S0002642"
                     /product="putative mitochondrial targeting signal"
     CDS             277. .2358
                     /gene="lys4"
                     /EC_number="4.2.1.36"
                     /codon_start=1
                     /db_xref="SGD:S0002642"
                     /product="homoaconitase"
                     /protein_id="CAA63764.1"
                     /db_xref="PID:e211801"
                     /db_xref="PID:g1089840"
                     /db_xref="GI:1089840"
                     /db_xref="SWISS-PROT:P49367"
                     /translation="MLRSTTFTRSFHSSRAWLKGQNLTEKIVQSYAVNLPEGKVVHSG
                     DYVSIKPAHCMSHDNSWPVALKFMGLGATKIKNPSQIVTTLDHDIQNKSEKNLTKYKN
                     IENFAKKHHIDHYPAGRGIGHQIMIEEGYAFPLNMTVASDSHSNTYGGLGSLGTPIVR
                     TDAAAIWATGQTWWQIPPVAQVELKGQLPQGVSGKDIIVALCGLFNNDQVLNHAIEFT
                     GDSLNALPIDHRLTIANMTTEWGALSGLFPVDKTLIDWYKNRLQKLGTNNHPRINPKT
                     IRALEEKAKIPKADKDAHYAKKLIIDLATLTHYVSGPNSVKVSNTVQDLSQQDIKINK
                     AYLVSCTNSRLSDLQSAADVVCPTGDLNKVNKVAPGVEFYVAAASSEIEADARKSGAW
                     EKLLKAGCIPLPSGCGPCIGLGAGLLEPGEVGISATNRNFKGRMGSKDALAYLASPAV
                     VAASAVLGKISSPAEVLSTSEIPFSGVKTEIIENPVVEEEVNAQTEAPKQSVEILEGF
                     PREFSGELVLCDADNINTDGIYPGKYTYQDDVPKEKMAQVCMENYDAEFRTKVHPGDI
                     VVSGFNFGTGSSREQAATALLAKGINLVVSGSFGNIFSRNSINNALLTLEIPALIKKL
                     REKYQGAPKELTRRTGWFLKWDVADAKVVVTEGSLDGPVILEQKVGELGKNLQEIIVK
                     GGLEGWVKSQL"
     gene            277. .2358
                     /gene="lys4"
                     /db_xref="SGD:S0002642"
     mat_peptide     337. .2355
                     /gene="lys4"
                     /db_xref="SGD:S0002642"
BASE COUNT      814 a    558 c    587 g    757 t
ORIGIN


  Query Match             67.8%;  Score 15.6;  DB 7;  Length 2716;
  Best Local Similarity   65.2%;  Pred. No. 3.4e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 gcngtngartggaayttygcnmg 23
          || |  ||:|||||:||:||  |
Db   1704 GCCGCTGAATGGAATTTCGCTTG 1682


RESULT   6
SC8419/c
LOCUS       SC8419       30507 bp    DNA             PLN       11-AUG-1997
DEFINITION  S.cerevisiae chromosome IV cosmid 8419.
ACCESSION   Z49701 Z71256
```

```
NID         g817819
VERSION     Z49701.1  GI:817819
KEYWORDS    AMDY; amidase; coatomer; dehydratase; delta element; gluconokinase;
            PAM1; protein kinase; PRP28; ribosomal protein; SEC26; sigma
            element; SUP2; transfer RNA-Tyr.
SOURCE      baker's yeast.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
            Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 30507)
  AUTHORS   Oliver,K. and Harris,D.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 30507)
  AUTHORS   Barrell,B. and Rajandream,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-1995) Saccharomyces cerevisiae chromosome IV
            sequencing project, Sanger Centre, Hinxton Hall, Hinxton, Cambridge
            CB10 1RQ E-mail: barrell@sanger.ac.uk
COMMENT     Notes:
            All CDS over 100 codons have been analysed. CDS that are completely
            overlapped and those that are overlapped by more than 50%
            of their length by a larger CDS have been omitted from this
            analysis.
            Details of the omitted CDS are available on request. The more
            significant matches with motifs in the PROSITE database are
            also included but some of these may be fortuitous. The length in
            codons and the calculated codon adaptation index (CAI)
            is given for each CDS.
            Cosmid 8419 is overlapped at the start by cosmid 9934, EMBL id
            SC9934,
            EMBL accession Z48612 and the end by cosmid 9320.
FEATURES             Location/Qualifiers
     source          1. .30507
                     /organism="Saccharomyces cerevisiae"
                     /strain="AB972"
                     /db_xref="taxon:4932"
                     /chromosome="IV"
                     /clone="cosmid 8419"
                     /map="4R"
     CDS             <1. .1128
                     /note="YD8419.01, putative dehydratase, incomplete orf -
                     overlaps orf YD9934.18, len: > 375, CAI: 0.23, similar to
                     3-isopropylmalate dehydratase and aconitate dehydratase"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="CAA89720.1"
                     /db_xref="PID:g817820"
                     /db_xref="GI:817820"
                     /translation="SNTVQDLSQQDIKINKAYLVSCTNSRLSDLQSAADVVCPTGDLN
                     KVNKVAPGVEFYVAAASSEIEADARKSGAWEKLLKAGCIPLPSGCGPCIGLGAGLLEP
                     GEVGISATNRNFKGRMGSKDALAYLASPAVVAASAVLGKISSPAEVLSTSEIPFSGVK
                     TEIIENPVVEEEVNAQTEAPKQSVEILEGFPREFSGELVLCDADNINTDGIYPGKYTY
                     QDDVPKEKMAQVCMENYDAEFRTKVHPGDIVVSGFNFGTGSSREQAATALLAKGINLV
                     VSGSFGNIFSRNSINNALLTLEIPALIKKLREKYQGAPKELTRRTGWFLKWDVADAKV
                     VVTEGSLDGPVILEQKVGELGKNLQEIIVKGGLEGWVKSQL"
     misc_feature    1. .19
                     /note="overlap with cosmid 9934, EMBL ID SC9934"
     CDS             1422. .3056
                     /note="YD8419.02, unknown, len: 544, CAI: 0.14"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="CAA89721.1"
                     /db_xref="PID:g817821"
                     /db_xref="GI:817821"
                     /db_xref="SPTREMBL:Q03776"
                     /translation="MDKYTALIHDENFSTLTLNVSRYPKSLAYWEKLLNYIVKASAPI
                     CKSTEPQLLKLIRCTYSSMLNEFPYLENYYIDFALLEYKLGNVSMSHKIFQRGLQAFN
                     QRSLLLWTSYLKFCNNVISHQKQLFKKYETAEEYVGLHFFSGEFWDLYLEQISSRCTS
                     SKKYWNVLRKILEIPLHSFSKFYALWLQRIDDIMDLKQLSQLTSKDELLKKLKIDINY
                     SGRKGPYLQDAKKKLKKITKEMYMVVQYQVLEIYSIFESKIYINYYTSPETLVSSDEI
                     ETWIKYLDYTITLQTDSLTHLNFQRALLPLAHYDLVWIKYSKWLINSKNDLLGAKNVL
                     LMGLKFSLKKTEIIKLLYSVICKLNEYVLLRNLLEKIESSYSDNVENVDDFEIFWDYL
                     QFKTFCQNSLYSSRYSDSQSNGLLNKELFDKVWKRLSCKEKKSGQEILLNNLVQFYSK
                     DTVEFVEKNIFQKIIEFGWEYYLQNGMFWNCYCRLIYFDTSRSYLDKRQYIVRKIWPQ
                     IDKKFAQSVLPSLTEFCESYFPEEMDTLEEMFTEEP"
     CDS             complement(3154. .3810)
                     /note="YD8419.03c, unknown, len: 218, CAI: 0.15"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="CAA89722.1"
                     /db_xref="PID:g817822"
                     /db_xref="GI:817822"
                     /db_xref="SPTREMBL:Q03778"
                     /translation="MFTWTIYVSLLLVLAGTFLMNRNTNTDIIDTFKREVDLPIPAQP
                     GPPFPLVTDYCDIVCGFGRGSAELGIPTANVPINQLPKGINDLDLGVYFGFAHIKTVD
                     GQELSVETRRDGRTVVYNYGQYLSEANDDLSVLPMVLSVGKNPFYGNDFKTMELHIIH
                     DFKNDFYGARVKFNILGHIRPELNYTTKEALIEDINIDIRTAQTVLATPPYQVFKQQL
                     "
     CDS             4533. .5411
                     /note="YD8419.04, len; 292, CAI: 0.16, C-terminal similar
                     to 50S ribosomal protein L5"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="CAA89723.1"
                     /db_xref="PID:g817823"
                     /db_xref="GI:817823"
                     /db_xref="SWISS-PROT:P36519"
                     /translation="MQRFSLVTHRSFSHSCVKPKSACSLVKPVHHLVKIDKSKLSPRF
                     PELKYDKSDIRSPGFKPKDTHADRLNDHYLNTLQSDLLLINYSHNAAVVKGLKQRAWS
                     GDSPYHLNRPPKNPRGSKAQLPDIHPIKWSNIPGLESVVINCFVREARENQLLAITAA
                     LQLQQITGCKPHPIFSKNDVPTWKLRKGHQMGAKVELKGKEMSQFLSTLTEIVLPRIR
                     EYKGISNQSGNRFGGISFGLTAEDIKFFPEIDANQDSWPKTFGMHININTSAQLDYQA
                     RTLLSGFQFPFFGEEK"
     gene            complement(5813. .8734)
                     /gene="SEC26"
                     /db_xref="SGD:S0002646"
     CDS             complement(5813. .8734)
                     /gene="SEC26"
                     /note="YD8419.05c, SEC26 gene, len: 973, CAI: 0.22,
                     EMBL:SC11236, U11236, coatomer beta subunit"
                     /codon_start=1
                     /db_xref="SGD:S0002646"
                     /product="Sec26p"
                     /protein_id="CAA89724.1"
                     /db_xref="PID:g817824"
                     /db_xref="GI:817824"
                     /db_xref="SWISS-PROT:P41810"
                     /translation="MTSLSSQPAYTLVFDPSPSMETYSSTDFQKALEKGSDEQKIDTM
                     KSILVTMLEGNPMPELLMHIIRFVMPSKNKELKKLLYFYWEIVPKLAEDGKLRHEMIL
                     VCNAIQHDLQHPNEYIRGNTLRFLTKLREAELLEQMVPSVLACLEYRHAYVRKYAILA
                     VFSIFKVSEHLLPDAKEIINSFIVAETDPICKRNAFIGLAELDRENALHYLENNIADI
                     ENLDPLLQAVFVQFIRQDANRTPALKAQYIELLMELLSTTTSDEVIFETALALTVLSA
                     NPNVLVPAVNKLIDLAVKVSDNNIKLIVLDRIQDINANNVGALEELTLDILRVLNAED
                     LDVRSKALDISMDLATSRNAEDVVQLLKKELQTTVNNPDQDKAMQYRQLLIKTIRTVA
                     VNFVEMAASVVSLLLDFIGDLNSVAASGIIAFIKEVIEKYPQLRANILENMVQTLDKV
                     RSAKAYRGALWIMGEYAEGESEIQHCWKHIRNSVGEVPILQSEIKKLTQNQEHTEENE
                     VDATAKPTGPVILPDGTYATESAFDVKTSQKSVTDEERDSRPPIRRFVLSGDFYTAAI
                     LANTIIKLVLKFENVSKNKTVINALKAEALLILVSIVRVGQSSLVEKKIDEDSLERVM
                     TSISILLDEVNPEEKKEEVKLLEVAFLDTTKSSFKRQIEIAKKNKHKRALKDSCKNIE
                     PIDTPISFRQFAGVDSTNVQKDSIEEDLQLAMKGDAIHATSSSSISKLKKIVPLCGFS
                     DPVYAEACITNNQFDVVLDVLLVNQTKETLKNLHVQFATLGDLKIIDTPQKTNVIPHG
                     FHKFTVTVKVSSADTGVIFGNIIYDGAHGEDARYVILNDVHVDIMDYIKPATADDEHF
                     RTMWNAFEWENKISVKSQLPTLHAYLRELVKGTNMGILTPSESLGEDDCRFLSCNLYA
                     KSSFGEDALANLCIEKDSKTNDVIGYVRIRSKGQGLALSLGDRVALIAKKTNKLALTH
                     V"
     CDS             complement(8975. .11338)
                     /note="YD8419.06c, unknown, len: 787, CAI: 0.12, potential
                     coiled-coil region"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="CAA89725.1"
                     /db_xref="PID:g817825"
                     /db_xref="GI:817825"
                     /db_xref="SPTREMBL:Q03780"
                     /translation="MFDGFSNNKGKRRSFFRFGSESKNNDSEKSVRKPSVPSTIKKST
                     NIARTSTAETSAPDIPPRSPNRNAHSRSHSIQAPLQKETLKNTNPFLNAEDTLGDSLE
```

```
                LTQSKEASGNDHKGIEYLQENNIIGQRTNPFTTSANSNAHFSKIKRSRPPPPPMDMKS
                ITTSISNNTTKEEIESNNDSERDSIAISSTHNQHRRQRSEAEKLVDDIENYINEHKVS
                SGSSLSLDTSENSDTKASQDKLPVDVMEAPILRNVSAESSLSYVKPLIVDNEEVNKAS
                NGNLVQSDHLKEFSSNLDDGDDKFSFSTSASGKSTKSLQQVSKDESSGFKAAHFDFAY
                KSNEHLGSDGSIASARKPLRITNEIDSGSSNEDDDDGLQEKGFVDSESKAFINYASDQ
                GSSIKNDVSTQEPELPSHRRIFRVVNEDRPSFYLNSVNDTGSLTDKHSFDTASSGEYD
                AKSNFSSQSGLSISKGSKSTVLAALDSNGNTKSSNKTSELNSLNSISESLVPAAHSFN
                EHTVTIPATVDLPNPVHDAPSERSVKCSPLTSVVSNKSEKSVPLVSSYVEELRLKYYK
                TSNFLQAPPNLPVALKQKNNLIQPKNIKVKLRTSSKQIGIKHGKVKQKLLALETRNEE
                SDGTATGLKNKINVDHTKEFHKLLGKENETGSISKKEGTDAEQAEDYLKDIPGDEAYN
                SDDIMAPLREKRGQNGSVDSVSRSNTVVSYYTRSQNRMRSGTLDNDYVNRQKLPTHIS
                LQDYRDANARSNISRQDSVSTTNSDVVDLSYSLGHGLRVANPDSDPE"
     CDS        complement(11592. .13070)
                /note="YD8419.07c, unknown, len: 492, CAI: 0.13"
                /codon_start=1
                /product="unknown"
                /protein_id="CAA89726.1"
                /db_xref="PID:g817826"
                /db_xref="GI:817826"
                /db_xref="SPTREMBL:Q03782"
                /translation="MRPRRRGLAYHHTKPKGQLSQGHYPTTSNDGQRRKVGNSEAFQS
                FDIWKNLDRIRSTKKNAGQFIKGSLLILPMRTEDKQQFDECMDELHKYISKDILRCYP
                QKEQKDEGMLFYIVLKDFNILDSCFVLSVLLAFQKRLWMAPSEKSYFRVPKNINLTGS
                FYLPKNIETGRGHIITSYRREQPSSSIVEVGFNVVPDFQQFQVKACHVSKFMNELSNF
                FSQVEFGKCEANVINYFKREYNRTYSQISLALYELPLIGDGLFDIKSYISKTRPIIET
                SKAQMIKHISEMKAYNEISGLQGDQFPRQQRPLSNSPSSNSISSSQTIEAGATSYQTQ
                PQRHAVNKPSNVLNSSNRHSGPKTFEDGRYSEGNKPGFMTQDEIKQHCIGTIKASMDA
                VKKKSSYQILKTYVRCPRQNYIDIVYQNLNDLRSKTNCNIVVLNLNNLHESQMWLESL
                NTTNYTIFAQAPHPSTIRVISIGGVGEYIVKALELILNILEH"
     CDS        13069. .13356
                /note="YD8419.08, questionable orf, overlaps delta
                element, len: 95, CAI: 0.10"
                /codon_start=1
                /product="unknown"
                /protein_id="CAA89727.1"
                /db_xref="PID:g817827"
                /db_xref="GI:817827"
                /db_xref="SPTREMBL:Q03783"
                /translation="MILIFFAYISRRETVSQRQTSTIVYNITLMFSTLRCALKILAQI
                KKRLQQTSSNHENKKYEEVIKDHVGVKIQNIICLLAFVLVVYYQMRCKKMT"
     LTR        13269. .13591
                /note="solo delta element"
     LTR        complement(13873. .14213)
                /note="solo sigma element"
     tRNA       14230. .14318
                /note="SUP2, tRNA-Tyr gene, len: 89, anticodon gta,
                potential intron between positions 14269 and 14282"
     gene       14725. .16374
                /gene="AMDY"
                /db_xref="SGD:S0002650"
     CDS        14725. .16374

  Query Match             67.8%;  Score 15.6;  DB 7;  Length 30507;
  Best Local Similarity   65.2%;  Pred. No. 2.7e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 gcngtngartggaayttygcnmg 23
           || |  ||:|||||:||:||  |
Db     474 GCCGCTGAATGGAATTTCGCTTG 452


RESULT   7
SCU46154/c
LOCUS       SCU46154     2710 bp    DNA             PLN       03-MAY-1999
DEFINITION  Sacchoromyces cerevisiae lysine biosynthesis enzyme homoaconitase
            (LYS4) gene, complete cds.
ACCESSION   U46154
NID         g1197580
VERSION     U46154.1  GI:1197580
KEYWORDS    .
SOURCE      baker's yeast.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
            Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (sites)
  AUTHORS   Wang,L., Okamoto,S. and Bhattacharjee,J.K.
  TITLE     Cloning and physical characterization of linked lysine genes (lys4,
            lys15) of Saccharomyces cerevisiae
  JOURNAL   Curr. Genet. 16 (1), 7-12 (1989)
  MEDLINE   90003361
REFERENCE   2  (bases 1 to 2710)
  AUTHORS   Irvin,S.D. and Bhattacharjee,J.K.
  TITLE     A unique fungal lysine biosynthesis enzyme shares a common ancestor
            with tricarboxylic acid cycle and leucine biosynthetic enzymes
            found in diverse organisms
  JOURNAL   J. Mol. Evol. 46 (4), 401-408 (1998)
  MEDLINE   98210230
REFERENCE   3  (bases 1 to 2710)
  AUTHORS   Irvin,S.D. and Bhattacharjee,J.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-1996) Jnanendra K. Bhattacharjee, Microbiology,
            Miami University, 32 Pearson Hall, Oxford, OH 45056, USA
FEATURES             Location/Qualifiers
     source          1. .2710
                     /organism="Saccharomyces cerevisiae"
                     /db_xref="taxon:4932"
                     /chromosome="IV"
                     /clone_lib="genomic library of S. Okamoto (1983), Abstr.
                     Meet. Mol. Biol. yeast, Cold Spring Harbor laboratory, NY,
                     p. 29"
     mRNA            <276. .>2357
                     /gene="LYS4"
                     /product="lysine biosynthesis enzyme homoaconitase"
                     /db_xref="SGD:S0002642"
     gene            <276. .>2357
                     /gene="LYS4"
                     /db_xref="SGD:S0002642"
     CDS             276. .2357
                     /gene="LYS4"
                     /EC_number="4.2.1.36"
                     /note="cis-homoaconitate hydratase"
                     /codon_start=1
                     /db_xref="SGD:S0002642"
                     /product="lysine biosynthesis enzyme homoaconitase"
                     /protein_id="AAA88902.1"
                     /db_xref="PID:g1197581"
                     /db_xref="GI:1197581"
                     /translation="MLRSTTFTRSFHSSRAWLKGQNLTEKIVQSYAVNLPEGKVVHSG
                     DYVSIKPAHCMSHDNSWPVALKFMGLGATKIKNPSQIVTTLDHDIQNKSEKNLTKYKN
                     IENFAKKHHIDHYPAGRGIGHQIMIEEGYAFPLNMTVASDSHSNTYGGLGSLGTPIVR
                     TDAAAIWATGQTWWQIPPVAQVELKGQLPQGVSGKDIIVALCGLFNNDQVLNHAIEFT
                     GDSLNALPIDHRLTIANMTTEWGALSGLFPVDKTLIDWYKNRLQKLGTNNHPRINPKT
                     IRALEEKAKIPKADKDAHYAKKLIIDLATLTHYVSGPNSVKVSNTVQDLSQQDIKINK
                     AYLVSCTNSRLSDLQSAADVVCPTGDLNKVNKVAPGVEFYVAAASSEIEADARKSGAW
                     EKLLKAGCIPLPSGCGPCIGLGAGLLEPGEVGISATNRNFKGRMGSKDALAYLASPAV
                     VAASAVLGKISSPAEVLSTSEIPFSGVKTEIIENPVVEEEVNAQTEAPKQSVEILEGF
                     PREFSGELVLCDADNINTDGIYPGKYTYQDDVPKEKMAQVCMENYDAEFRTKVHPGDI
                     VVSGFNFGTGSSREQAATALLAKGINLVVSGSFGNIFSRNSINNALLTLEIPALIKKL
                     REKYQGAPKELTRRTGWFLKWDVADAKVVVTEGSLDGPVILEQKVGELGKNLQEIIVK
                     GGLEGWVKSQL"
BASE COUNT      813 a    556 c    585 g    756 t
ORIGIN


  Query Match             67.8%;  Score 15.6;  DB 8;  Length 2710;
  Best Local Similarity   65.2%;  Pred. No. 3.4e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 gcngtngartggaayttygcnmg 23
           || |  ||:|||||:||:||  |
Db    1703 GCCGCTGAATGGAATTTCGCTTG 1681


RESULT   8
HSAJ6276
LOCUS       HSAJ6276     4564 bp    mRNA            PRI       08-FEB-1999
```

```
DEFINITION  Homo sapiens mRNA for transient receptor potential protein TRP6.
ACCESSION   AJ006276
NID         g4454260
VERSION     AJ006276.1  GI:4454260
KEYWORDS    transient receptor potential protein; trp6 gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4564)
  AUTHORS   D'Esposito,M., Strazzullo,M., Cuccurese,M., Spalluto,C., Rocchi,M.,
            D'Urso,M. and Ciccodicola,A.
  TITLE     Identification and assignment of the human transient receptor
            potential channel 6 gene TRPC6 to chromosome 11q21-->q22
  JOURNAL   Cytogenet. Cell Genet. 83 (1-2), 46-47 (1998)
  MEDLINE   99126323
REFERENCE   2  (bases 1 to 4564)
  AUTHORS   D'Esposito,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-1998) International Institute of Genetics and
            Biophysics, Via Marconi, 10, Napoli 80125, ITALY
FEATURES             Location/Qualifiers
     source          1. .4564
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="q21-22"
     gene            428. .3223
                     /gene="TRP6"
     CDS             428. .3223
                     /gene="TRP6"
                     /codon_start=1
                     /evidence=experimental
                     /product="transient receptor potential protein"
                     /protein_id="CAA06943.1"
                     /db_xref="PID:e1386893"
                     /db_xref="PID:g4454261"
                     /db_xref="GI:4454261"
                     /translation="MSQSPAFGPRRGSSPRGAAGAAARRNESQDYLLMDSELGEDGCP
                     QAPLPCYGYYPCFRGSDNRLAHRRQTVLREKGRRLANRGPAYMFSDRSTSLSIEEERF
                     LDAAEYGNIPVVRKMLEECHSLNVNCVDYMGQNALQLAVANEHLEITELLLKKENLSR
                     VGDALLLAISKGYVRIVEAILSHPAFAEGKRLATSPSQSELQQDDFYAYDEDGTRFSH
                     DVTPIILAAHCQEYEIVHTLLRKGARIERPHDYFCKCNDCNQKQKHDSFSHSRSRINA
                     YKGLASPAYLSLSSEDPVMTALELSNELAVLANIEKEFKNDYKKLSMQCKDFVVGLLD
                     LCRNTEEVEAILNGDVETLQSGDHGRPNLSRLKLAIKYEVKKFVAHPNCQQQLLSIWY
                     ENLSGLRQQTMAVKFLVVLAVAIGLPFLALIYWFAPCSKMGKIMRGPFMKFVAHAASF
                     TIFLGLLVMNAADRFEGTKLLPNETSTDNAKQLFRMKTSCFSWMEMLIISWVIGMIWA
                     ECKEIWTQGPKEYLFELWNMLDFGMLAIFAASFIARFMAFWHASKAQSIIDANDTLKD
                     LTKVTLGDNVKYYNLARIKWDPSDPQIISEGLYAIAVVLSFSRIAYILPANESFGPLQ
                     ISLGRTVKDIFKFMVIFIMVFVAFMIGMFNLYSYYIGAKQNEAFTTVEESFKTLFWAI
                     FGLSEVKSVVINYNHKFIENIGYVLYGVYNVTMVIVLLNMLIAMINSSFQEIEDDADV
                     EWKFARAKLWFSYFEEGRTLPVPFNLVPSPKSLFYLLLKLKKWISELFQGHKKGFQED
                     AEMNKINEEKKLGILGSHEDLSKLSLDKKQVGHNKQPSIRSSEDFHLNSFNNPPRQYQ
                     KIMKRLIKRYVLQAQIDKESDEVNEGELKEIKQDISSLRYELLEEKSQNTEDLAELIR
                     ELGEKLSMEPNQEETNR"
BASE COUNT     1283 a    990 c   1021 g   1270 t
ORIGIN


  Query Match             67.8%;  Score 15.6;  DB 10;  Length 4564;
  Best Local Similarity   65.2%;  Pred. No. 3.2e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 gcngtngartggaayttygcnmg 23
           |  || ||:|||||  ||:|| :|
Db    2642 GATGTGGAGTGGAAATTTGCAAG 2664


RESULT  9
HSTRPC3GN
LOCUS       HSTRPC3GN     353 bp    mRNA            PRI       06-JUN-1996
DEFINITION  H.sapiens mRNA for TRPC3 protein.
ACCESSION   X89068
NID         g1019789
VERSION     X89068.1  GI:1019789
KEYWORDS    TRPC3 gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 353)
  AUTHORS   Wes,P.D., Chevesich,J., Jeromin,A., Rosenberg,C., Stetten,G. and
            Montell,C.
  TITLE     TRPC1, a human homolog of a Drosophila store-operated channel
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 92 (21), 9652-9656 (1995)
  MEDLINE   96003837
REFERENCE   2  (bases 1 to 353)
  AUTHORS   Montell,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-1995) C. Montell, The Johns Hopkins Univ. School
            of Med., Dept of Biological Chemistry, 725 N. Wolfe Street,
            Baltimore MD 21218, USA
COMMENT     Related sequence R34716.
FEATURES             Location/Qualifiers
     source          1. .353
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /dev_stage="fetal"
                     /tissue_type="brain"
     gene            1. .353
                     /gene="trpc3"
     CDS             <1. .>353
                     /gene="trpc3"
                     /note="transient receptor potential channel related"
                     /codon_start=3
                     /product="TRPC3 protein"
                     /protein_id="CAA61448.1"
                     /db_xref="PID:g1019790"
                     /db_xref="GI:1019790"
                     /db_xref="SPTREMBL:Q15660"
                     /translation="TSVVLKYDHKFIENIGYVLYGIYNVTMVVVLLNMLIAMINSSYQ
                     EIEDDSDVEWKFARSKLWLSYFDDGKTLPPPFSLVPSPKSFVYFIMRIVNFPKCRRRR
                     LQKDIGNGEWGNSKS"
BASE COUNT      113 a     56 c     69 g    115 t
ORIGIN


  Query Match             67.8%;  Score 15.6;  DB 10;  Length 353;
  Best Local Similarity   65.2%;  Pred. No. 4.2e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 gcngtngartggaayttygcnmg 23
           |  || ||:|||||  ||:|| :|
Db     153 GATGTAGAATGGAAGTTTGCTCG 175


RESULT  10
HSY13758
LOCUS       HSY13758     3448 bp    mRNA            PRI       01-JUL-1997
DEFINITION  Homo sapiens mRNA for transient receptor potential related channel
            3 protein.
ACCESSION   Y13758
NID         g2225936
VERSION     Y13758.1  GI:2225936
KEYWORDS    transient receptor potential channel; TRPC3 gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3448)
  AUTHORS   Xu,X.Z., Li,H.S., Guggino,W.B. and Montell,C.
  TITLE     Coassembly of TRP and TRPL produces a distinct store-operated
            conductance
  JOURNAL   Cell 89 (7), 1155-1164 (1997)
  MEDLINE   97358541
REFERENCE   2  (bases 1 to 3448)
```

```
  AUTHORS   Montell,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1997) C. Montell, The Johns Hopkins University
            School of Medicine, Department of Biological Chermistry, 725 N.
            Wolfe Street, Baltimore, MD 21205, USA
COMMENT     Related sequences: X89068, U47050.
FEATURES             Location/Qualifiers
     source          1. .3448
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            425. .2971
                     /gene="TRPC3"
     CDS             425. .2971
                     /gene="TRPC3"
                     /codon_start=1
                     /product="transient receptor potential related channel 3
                     protein"
                     /protein_id="CAA74083.1"
                     /db_xref="PID:e323535"
                     /db_xref="PID:g2225937"
                     /db_xref="GI:2225937"
                     /db_xref="SPTREMBL:Q13507"
                     /translation="MEGSPSLRRMTVMREKGRRQAVRGPAFMFNDRGTSLTAEEERFL
                     DAAEYGNIPVVRKMLEESKTLNVNCVDYMGQNALQLAVGNEHLEVTELLLKKENLARI
                     GDALLLAISKGYVRIVEAILNHPGFAASKRLTLSPCEQELQDDDFYAYDEDGTRFSPD
                     ITPIILAAHCQKYEVVHMLLMKGARIERPHDYFCKCGDCMEKQRHDSFSHSRSRINAY
                     KGLASPAYLSLSSEDPVLTALELSNELAKLANIEKEFKNDYRKLSMQCKDFVVGVLDL
                     CRDSEEVEAILNGDLESAEPLEVHRHKASLSRVKLAIKYEVKKFVAHPNCQQQLLTIW
                     YENLSGLREQTIAIKCLVVLVVALGLPFLAIGYWIAPCSRLGKILRSPFMKFVAHAAS
                     FIIFLGLLVFNASDRFEGITTLPNITVTDYPKQIFRVKTTQFTWTEMLIMVWVLGMMW
                     SECKELWLEGPREYILQLWNVLDFGMLSIFIAAFTARFLAFLQATKAQQYVDSYVQES
                     DLSEVTLPPEIQYFTYARDKWLPSDPQIISEGLYAIAVVLSFSRIAYILPANESFGPL
                     QISLGRTVKDIFKFMVLFIMVFFAFMIGMFILYSYYLGAKVNAAFTTVEESFKTLFWS
                     IFGLSEVTSVVLKYDHKFIENIGYVLYGIYNVTMVVVLLNMLIAMINSSYQEIEDDSD
                     VEWKFARSKLWLSYFDDGKTLPPPFSLVPSPKSFVYFIMRIVNFPKCRRRRLQKDIEM
                     GMGNSKSRLNLFTQSNSRVFESHSFNSILNQPTRYQQIMKRLIKRYVLKAQVDKENDE
                     VNEGELKEIKQDISSLRYELLEDKSQATEELAILIHKLSEKLNPSMLRCE"
BASE COUNT      945 a    772 c    802 g    929 t
ORIGIN


  Query Match             67.8%;  Score 15.6;  DB 10;  Length 3448;
  Best Local Similarity   65.2%;  Pred. No. 3.3e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 gcngtngartggaayttygcnmg 23
          |  || ||:||||| ||:|| :|
Db   2468 GATGTAGAATGGAAGTTTGCTCG 2490


RESULT  11
AC005154
LOCUS       AC005154   151630 bp    DNA             PRI       23-SEP-1998
DEFINITION  Homo sapiens PAC clone DJ0777O23 from 7p14-p15, complete sequence.
ACCESSION   AC005154
NID         g3242763
VERSION     AC005154.1  GI:3242763
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 151630)
  AUTHORS   Bauer,C., Langston,Y. and Harrison,M.
  TITLE     The sequence of Homo sapiens PAC clone DJ0777O23
  JOURNAL   Unpublished (1998)
REFERENCE   2  (bases 1 to 151630)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     SUBMITTED BY:  WUGSC
            Genome Sequencing Center
            Department of Genetics
            Washington University
            St. Louis MO 63108, USA
            http://genome.wustl.edu/gsc
            mailto:sapiens@watson.wustl.edu

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded or sequenced with an alternate
            chemistry; an attempt was made to resolve all sequencing problems,
            such as compressions and repeats; all regions were covered by
            sequence from more than one subclone; and the assembly was
            confirmed by restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center.  For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            This clone was derived from human PAC library RPCI-4, prepared by
            Pieter de Jong and coworkers at Roswell Park Cancer Institute,
            using the method described by Ioannou et al., Nature Genetics
            6:84-9 (1994).  The library is from one male donor.  For further
            details, see http://bacpac.med.buffalo.edu/
            The clone is available from Genome Systems, Inc.
            (http://www.genomesystems.com).
            VECTOR:  pCYPAC2
            NEIGHBORING SEQUENCE INFORMATION:
            The actual start of this clone is at base position 1 of DJ0777O23;
            actual end is at 151630 of DJ0777O23.
FEATURES             Location/Qualifiers
     source          1. .151630
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7p14-p15"
                     /clone_lib="RPCI-4"
                     /clone="DJ0777O23"
     repeat_region   1. .57
                     /rpt_family="Alu"
     repeat_region   1335. .1634
                     /rpt_family="Alu"
     repeat_region   3612. .3770
                     /rpt_family="MER1_type"
     repeat_region   3771. .4076
                     /rpt_family="Alu"
     repeat_region   4078. .4238
                     /rpt_family="MER1_type"
     repeat_region   4336. .4636
                     /rpt_family="Alu"
     repeat_region   4821. .4933
                     /rpt_family="L1"
     repeat_region   5121. .5256
                     /rpt_family="L1"
     repeat_region   5266. .5566
                     /rpt_family="Alu"
     repeat_region   5597. .5714
                     /rpt_family="L1"
     repeat_region   5788. .5861
                     /rpt_family="L1"
     repeat_region   5873. .6007
                     /rpt_family="L1"
     repeat_region   6988. .7049
```

```
                /rpt_family="L2"
repeat_region   7357. .7481
                /rpt_family="MIR"
repeat_region   7512. .7628
                /rpt_family="MIR"
STS             8266. .8530
                /db_xref="GI:1343978"
repeat_region   9111. .9287
                /rpt_family="Alu"
repeat_region   10298. .10600
                /rpt_family="Alu"
repeat_region   12676. .12895
                /rpt_family="MIR"
repeat_region   13006. .13047
                /rpt_family="L2"
repeat_region   14748. .14827
                /rpt_family="L2"
repeat_region   15146. .15465
                /rpt_family="MaLR"
repeat_region   15488. .15595
                /rpt_family="L2"
repeat_region   16449. .16629
                /rpt_family="MIR"
repeat_region   16665. .16815
                /rpt_family="MIR"
repeat_region   16844. .17000
                /rpt_family="MIR"
repeat_region   17006. .17349
                /rpt_family="MER1_type"
repeat_region   19648. .19739
                /rpt_family="MIR"
repeat_region   20111. .20419
                /rpt_family="Alu"
repeat_region   20804. .20970
                /rpt_family="MER1_type"
repeat_region   22168. .22551
                /rpt_family="Retroviral"
repeat_region   22552. .22979
                /rpt_family="Retroviral"
repeat_region   23039. .25680
                /rpt_family="Retroviral"
repeat_region   25674. .25855
                /rpt_family="Retroviral"
repeat_region   25924. .27079
                /rpt_family="Retroviral"
repeat_region   27097. .27517
                /rpt_family="Retroviral"
repeat_region   27520. .27882
                /rpt_family="Retroviral"
repeat_region   27910. .28343
                /rpt_family="MER4-group"
repeat_region   28565. .29076
                /rpt_family="MaLR"
repeat_region   29241. .29379
                /rpt_family="L1"
repeat_region   29448. .29567
                /rpt_family="(GAAA)n"
repeat_region   29605. .29913
                /rpt_family="Alu"
repeat_region   29920. .30328
                /rpt_family="L1"
repeat_region   30337. .31163
                /rpt_family="L1"
repeat_region   31167. .31347
                /rpt_family="Alu"
repeat_region   31348. .32244
                /rpt_family="L1"
repeat_region   32241. .32867
                /rpt_family="L1"
repeat_region   32882. .33181
                /rpt_family="L1"
repeat_region   33184. .33479
                /rpt_family="Alu"
repeat_region   33481. .33965
                /rpt_family="L1"
repeat_region   33973. .35027
                /rpt_family="L1"
repeat_region   35127. .36623
                /rpt_family="L1"
repeat_region   36663. .36960
                /rpt_family="Alu"
repeat_region   37383. .37472
                /rpt_family="MER86"
repeat_region   37602. .37898
                /rpt_family="Alu"
repeat_region   38596. .38895
                /rpt_family="Alu"
repeat_region   38947. .39127
                /rpt_family="MIR"
STS             39222. .39481
                /db_xref="GI:859022"
gene            complement(39691. .47308)
                /gene="WUGSC:H_DJ0777O23.1"
CDS             complement(join(39691. .39834,41402. .41537,43135. .43280,
                47168. .47308))
                /gene="WUGSC:H_DJ0777O23.1"
                /note="similar to protein U28928 (PID:g861306);
                H_DJ0777O23.1"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC23790.1"
                /db_xref="PID:g3242764"
                /db_xref="GI:3242764"
                /translation="MANSGCKDVTGPDEESFLYFAYGSNLLTERIHLRNPSAAFFCVA
                RLQDFKLDFGNSQGKTSQTWHGGIATIFQSPGDEVWGVVWKMNKSNLNSLDEQEGVKS
                GMYVVIEVKVATQEGKEITCRSYLMTNYESAPPSPQYKKIICMGAKENGLPLEYQEKL

  Query Match             67.8%;  Score 15.6;  DB 11;  Length 151630;
  Best Local Similarity   65.2%;  Pred. No. 2.3e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 gcngtngartggaayttygcnmg 23
          |  || ||:||||| ||:|| :|
Db  86858 GATGTGGAGTGGAAATTTGCAAG 86880


RESULT  12
AF093117
LOCUS       AF093117   147216 bp    DNA             PRI       02-OCT-1998
DEFINITION  Homo sapiens chromosome 7qtelo BAC E3, complete sequence.
ACCESSION   AF093117
NID         g3687910
VERSION     AF093117.1  GI:3687910
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 147216)
  AUTHORS   Blechschmidt,K., Nordsiek,G., Drescher,B., Weber,J., Schattevoy,R.,
            Korenberg,J. and Rosenthal,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-SEP-1998) Genome Analysis, Institute for Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES             Location/Qualifiers
     source          1. .147216
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7qtelo"
                     /clone="BAC E3"
     repeat_region   complement(427. .480)
                     /rpt_family="MIR"
                     /evidence=not_experimental
     repeat_region   complement(481. .739)
                     /rpt_family="AluSg1"
```

```
                /evidence=not_experimental
repeat_region   complement(740. .884)
                /rpt_family="MIR"
                /evidence=not_experimental
repeat_region   complement(1701. .2121)
                /rpt_family="MSTA"
                /evidence=not_experimental
misc_feature    4532. .4669
                /note="CpG island score = 0.67, GC = 72.50%, CpGs = 6"
                /note="Region: CpG island"
                /evidence=not_experimental
misc_feature    4532. .4723
                /note="CpG island score = 0.62, GC = 72.90%, CpGs = 8"
                /note="Region: CpG island"
                /evidence=not_experimental
repeat_region   6845. .7072
                /rpt_family="L2"
                /evidence=not_experimental
repeat_region   7073. .7278
                /rpt_family="MER20"
                /evidence=not_experimental
repeat_region   7279. .7364
                /rpt_family="L2"
                /evidence=not_experimental
repeat_region   complement(8141. .8338)
                /rpt_family="L1MA9"
                /evidence=not_experimental
repeat_region   complement(8382. .8662)
                /rpt_family="AluSx"
                /evidence=not_experimental
repeat_region   complement(8663. .8939)
                /rpt_family="L1MA9"
                /evidence=not_experimental
repeat_region   complement(8940. .9249)
                /rpt_family="AluYb8"
                /evidence=not_experimental
repeat_region   complement(9250. .9877)
                /rpt_family="L1MA9"
                /evidence=not_experimental
repeat_region   10866. .11139
                /rpt_family="MER33"
                /evidence=not_experimental
repeat_region   11140. .11449
                /rpt_family="AluSx"
                /evidence=not_experimental
repeat_region   11450. .11497
                /rpt_family="MER33"
                /evidence=not_experimental
repeat_region   complement(11799. .11862)
                /rpt_family="L1ME"
                /evidence=not_experimental
repeat_region   complement(11863. .12173)
                /rpt_family="AluSq"
                /evidence=not_experimental
repeat_region   complement(12174. .12614)
                /rpt_family="L1ME"
                /evidence=not_experimental
repeat_region   12875. .13039
                /rpt_family="MER63A"
                /evidence=not_experimental
repeat_region   13640. .14178
                /rpt_family="MER69"
                /evidence=not_experimental
repeat_region   14902. .15137
                /rpt_family="MER69"
                /evidence=not_experimental
repeat_region   16032. .16356
                /rpt_family="L1MB8"
                /evidence=not_experimental
repeat_region   complement(16909. .17377)
                /rpt_family="Tigger3b"
                /evidence=not_experimental
repeat_region   complement(17378. .17688)
                /rpt_family="AluY"
                /evidence=not_experimental
repeat_region   complement(17689. .18156)
                /rpt_family="Tigger3b"
                /evidence=not_experimental
repeat_region   complement(18159. .18326)
                /rpt_family="Tigger3b"
                /evidence=not_experimental
repeat_region   18468. .18817
                /rpt_family="THE1B"
                /evidence=not_experimental
repeat_region   19803. .19918
                /rpt_family="LTR33"
                /evidence=not_experimental
repeat_region   complement(21579. .21707)
                /rpt_family="L2"
                /evidence=not_experimental
repeat_region   complement(22736. .22773)
                /rpt_family="MIR"
                /evidence=not_experimental
repeat_region   complement(22868. .23152)
                /rpt_family="AluSx"
                /evidence=not_experimental
repeat_region   complement(23232. .23508)
                /rpt_family="AluJo"
                /evidence=not_experimental
repeat_region   complement(23594. .23625)
                /rpt_family="MIR"
                /evidence=not_experimental
misc_feature    29361. .29464
                /note="CpG island score = 1.18, GC = 69.20%, CpGs = 13"
                /note="Region: CpG island"
                /evidence=not_experimental
repeat_region   complement(29487. .29596)
                /rpt_family="L2"
                /evidence=not_experimental
misc_feature    29788. .29913
                /note="CpG island score = 0.93, GC = 66.70%, CpGs = 14"
                /note="Region: CpG island"
                /evidence=not_experimental
misc_feature    30121. .30366
                /note="CpG island score = 0.76, GC = 66.30%, CpGs = 19"
                /note="Region: CpG island"
                /evidence=not_experimental
repeat_region   31204. .31240
                /rpt_family="MER91C"
                /evidence=not_experimental
repeat_region   complement(31487. .31753)
                /rpt_family="AluJo"
                /evidence=not_experimental
repeat_region   32049. .32358
                /rpt_family="AluSq"
                /evidence=not_experimental
repeat_region   32507. .32686
                /rpt_family="MLT1G"
                /evidence=not_experimental
repeat_region   complement(32914. .33550)
                /rpt_family="HERV16"
                /evidence=not_experimental
repeat_region   complement(33551. .33892)
                /rpt_family="MER7A"
                /evidence=not_experimental
repeat_region   complement(33893. .34123)
                /rpt_family="HERV16"
                /evidence=not_experimental
repeat_region   34291. .34774
                /rpt_family="MLT2CB"
                /evidence=not_experimental
repeat_region   complement(34919. .35067)
                /rpt_family="Tigger1"
                /evidence=not_experimental
repeat_region   complement(35068. .35320)
                /rpt_family="AluY"
```

```
                     /evidence=not_experimental
     repeat_region   complement(35321. .35397)
                     /rpt_family="Tigger1"
                     /evidence=not_experimental
     repeat_region   35398. .35679
                     /rpt_family="AluSp"
                     /evidence=not_experimental
     repeat_region   complement(35714. .37038)
                     /rpt_family="Tigger1"
                     /evidence=not_experimental
     repeat_region   complement(37095. .37903)
                     /rpt_family="Tigger1"
                     /evidence=not_experimental
     repeat_region   complement(38696. .38845)
                     /rpt_family="HERV16"
                     /evidence=not_experimental
     repeat_region   39153. .39462
                     /rpt_family="AluSx"
                     /evidence=not_experimental
     repeat_region   39471. .39655
                     /rpt_family="MIR"
                     /evidence=not_experimental
     repeat_region   39688. .40142
                     /rpt_family="L1"
                     /evidence=not_experimental
     repeat_region   40146. .40573

  Query Match             67.8%;  Score 15.6;  DB 11;  Length 147216;
  Best Local Similarity   65.2%;  Pred. No. 2.3e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 gcngtngartggaayttygcnmg 23
          || || | :|||||:||:||  |
Db 107284 GCAGTTGCGTGGAACTTTGCAGG 107306


RESULT  13
HSU47050
LOCUS       HSU47050     3417 bp    mRNA            PRI       04-AUG-1997
DEFINITION  Human putative calcium influx channel (htrp3) mRNA, complete cds.
ACCESSION   U47050
NID         g2295902
VERSION     U47050.1  GI:2295902
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3417)
  AUTHORS   Zhu,X., Jiang,M., Peyton,M., Boulay,G., Hurst,R., Stefani,E. and
            Birnbaumer,L.
  TITLE     trp, a novel mammalian gene family essential for agonist-activated
            capacitative Ca2+ entry
  JOURNAL   Cell 85 (5), 661-671 (1996)
  MEDLINE   96234226
REFERENCE   2  (bases 1 to 3417)
  AUTHORS   Zhu,X., Peyton,M. and Birnbaumer,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JAN-1996) Xi Zhu, Anesthesiology, UCLA School of
            Medicine, BH-612, CHS, Los Angeles, CA 90095-1778, USA
COMMENT     On Aug 4, 1997 this sequence version replaced gi:1326112.
FEATURES             Location/Qualifiers
     source          1. .3417
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .3417
                     /gene="htrp3"
     CDS             163. .2709
                     /gene="htrp3"
                     /note="Htrp-3; putative calcium influx channel; stimulated
                     by agonist to some G protein coupled receptors;  similar
                     to Drosophila transient receptor potential"
                     /codon_start=1
                     /product="calcium influx channel"
                     /protein_id="AAC51653.1"
                     /db_xref="PID:g2295903"
                     /db_xref="GI:2295903"
                     /translation="MEGSPSLRRMTVMREKGRRQAVRGPAFMFNDRGTSLTAEEERFL
                     DAAEYGNIPVVRKMLEESKTLNVNCVDYMGQNALQLAVGNEHLEVTELLLKKENLARI
                     GDALLLAISKGYVRIVEAILNHPGFAASKRLTLSPCEQELQDDDFYAYDEDGTRFSPD
                     ITPIILAAHCQKYEVVHMLLMKGARIERPHDYFCKCGDCMEKQRHDSFSHSRSRINAY
                     KGLASPAYLSLSSEDPVLTALELSNELAKLANIEKEFKNDYRKLSMQCKDFVVGVLDL
                     CRDSEEVEAILNGDLESAEPLEVHRHKASLSRVKLAIKYEVKKFVAHPNCQQQLLTIW
                     YENLSGLREQTIAIKCLVVLVVALGLPFLAIGYWIAPCSRLGKILRSPFMKFVAHAAS
                     FIIFLGLLVFNASDRFEGITTLPNITVTDYPKQIFRVKTTQFTWTEMLIMVWVLGMMW
                     SECKELWLEGPREYILQLWNVLDFGMLSIFIAAFTARFLAFLQATKAQQYVDSYVQES
                     DLSEVTLPPEIQYFTYARDKWLPSDPQIISEGLYAIAVVLSFSRIAYILPANESFGPL
                     QISLGRTVKDIFKFMVLFIMVFFAFMIGMFILYSYYLGAKVNAAFTTVEESFKTLFWS
                     IFGLSEVTSVVLKYDHKFIENIGYVLYGIYNVTMVVVLLNMLIAMINSSYQEIEDDSD
                     VEWKFARSKLWLSYFDDGKTLPPPFSLVPSPKSFVYFIMRIVNFPKCRRRRLQKDIEM
                     GMGNSKSRLNLFTQSNSRVFESHSFNSILNQPTRYQQIMKRLIKRYVLKAQVDKENDE
                     VNEGELKEIKQDISSLRYELLEDKSQATEELAILIHKLSEKLNPSMLRCE"
BASE COUNT      965 a    744 c    769 g    939 t
ORIGIN


  Query Match             67.8%;  Score 15.6;  DB 11;  Length 3417;
  Best Local Similarity   65.2%;  Pred. No. 3.3e+02;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 gcngtngartggaayttygcnmg 23
          |  || ||:||||| ||:|| :|
Db   2206 GATGTAGAATGGAAGTTTGCTCG 2228


RESULT  14
AF057748
LOCUS       AF057748     2842 bp    mRNA            ROD       19-OCT-1998
DEFINITION  Mus musculus putative capacitative calcium entry channel (Trp6)
            mRNA, complete cds.
ACCESSION   AF057748
NID         g3766190
VERSION     AF057748.1  GI:3766190
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2842)
  AUTHORS   Buess,M., Engler,O., Hirsch,H.H. and Moroni,C.
  TITLE     Search for oncogenic regulators in an autocrine tumor model using
            differential display PCR: Identification of novel candidate genes
            including the calcium channel mtrp 6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2842)
  AUTHORS   Buess,M. and Moroni,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-APR-1998) Institute for Medical Microbiology,
            Petersplatz 10, Basel 4003, Switzerland
FEATURES             Location/Qualifiers
     source          1. .2842
                     /organism="Mus musculus"
                     /strain="DBA/2"
                     /db_xref="taxon:10090"
                     /cell_line="PB-3c"
     gene            1. .2842
                     /gene="Trp6"
     CDS             22. .2652
                     /gene="Trp6"
                     /codon_start=1
                     /product="putative capacitative calcium entry channel"
                     /protein_id="AAC64394.1"
                     /db_xref="PID:g3766191"
                     /db_xref="GI:3766191"
                     /translation="MSRGNENRLTHRRQTILREKGRRLANRGPAYMFNDHSTSLSIEE
                     ERFLDAAEYGNIPVVRKMLEECHSLNVNCVDYMGQDALQLAVANEHLEITELLLKKEN
```

LSRVGDALLLAISKGYVRIVEAILNHPAFAEGKRLATSPSQSELQQDDFYAYDEDGTR
FSHDVTPIILAAHCQEYEIVHTLLRKGARIERPHDYFCKCTECSQKQKHDSFSHSRSR
INAYKGLASPAYLSLSSEDPVMTALELSNELAVLANIEKEFKNDYRKLSMQCKDFVVG
LLDLCRNTEEVEAILNGDAETRQPGDFGRPNLSRLKLAIKYEVKKFVAHPNCQQQLLS
IWYENLSGLRQQTMAVKFLVVLAVAIGLPFLALIYWCAPCSKMGKILRGPFMKFVAHA
ASFTIFLGLLVMNAADRFEGTKLLPNETSTDNARQLFRMKTSCFSWMEMLIISWVIGM
IWAECKEIWTQGPKEYLFELWNMLDFGMLAIFAASFIARFMAFWHASKAQSIIDANDT
LKDLTKVTLGDNVKYYNLARIKWDPTDPQIISEGLYAIAVVLSFSRIAYILPANESFG
PLQISLGRTVKDIFKFMVIFIMVFVAFMIGMFNLYSYYIGAKQNEAFTTVEESFKTLF
WAIFGLSEVKSVVINYNHKFIENIGYVLYGVYNVTMVIVLLNMLIAMINSSFQEIEDD
ADVEWKFARAKLWFSYFEEGRTLPVPFNLVPSPKSLLYLLLKFKKWMCELIQGQKQGF
QEDAEMNKRNEEKKFGISGSHEDLSKFSLDKNQLAHNKQSSTRSSEDYHLNSFSNPPR
QYQKIMKRLIKRYVLQAQIDKESDEVNEGELKEIKQDISSLRYELLEEKSQNTEDLAE
LIRKLGERLSLEPKLEESRR"

BASE COUNT 873 a 565 c 654 g 750 t
ORIGIN

Query Match 67.8%; Score 15.6; DB 12; Length 2842;
Best Local Similarity 65.2%; Pred. No. 3.4e+02;
Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 gcngtngartggaayttygcnmg 23
| || ||:||||| ||:|| :|
Db 2071 GACGTGGAGTGGAAGTTTGCAAG 2093

RESULT 15
AF111107
LOCUS AF111107 3691 bp mRNA ROD 15-MAR-1999
DEFINITION Mus musculus transient receptor potential 2 (Trp2) mRNA, complete cds.
ACCESSION AF111107
NID g4324937
VERSION AF111107.1 GI:4324937
KEYWORDS .
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3691)
AUTHORS Vannier,B., Peyton,M., Boulay,G., Brown,D., Qin,N., Jiang,M., Zhu,X. and Birnbaumer,L.
TITLE Mouse trp2, the homologue of the human trpc2 pseudogene, encodes mTrp2, a store depletion-activated capacitative Ca2+ entry channel
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (5), 2060-2064 (1999)
MEDLINE 99162557
REFERENCE 2 (bases 1 to 3691)
AUTHORS Vannier,B., Peyton,M., Zhu,X. and Birnbaumer,L.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-1998) Anesthesiology, UCLA, 10833 Le Conte Avenue, Los Angeles, CA 90095-1778, USA
FEATURES Location/Qualifiers
source 1. .3691
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="testis"
gene 1. .3691
/gene="Trp2"
CDS 390. .3608
/gene="Trp2"
/note="putative capacitative calcium entry channel"
/codon_start=1
/product="transient receptor potential 2"
/protein_id="AAD17195.1"
/db_xref="PID:g4324938"
/db_xref="GI:4324938"
/translation="MGTKTHPVVPWSTKEISELKGMLKQLQPGPLGRAARMVLSAARK
APPASVVSPNNSHGEPGPSRAESAEPRAEEPNRKTAVGRRKRRKVQEPRRSLSNSSSQ
PNRRTGRTRQRQHRPQTKSDDGGVQAAGQCPICAGFFSIETLPQHAATCGESPPPQPA
SPASLSSSESVLRRHHVALTPVPLVPKPQPNWTEIVNKKLKFPPTLLRAIQEGQLGLV
QQLLESSSDASGAGPGGPLRNVEESEDRSWREALNLAIRLGHEVITDVLLANVKFDFR
QIHEALLVAVDTNQPAVVRRLLARLEREKGRKVDTKSFSLAFFDSSIDGSRFAPGVTP
LTLACQKDLYEIAQLLMDQGHTIARPHPVSCACLECSNARRYDLLKFSLSRINTYRGI
ASRAHLSLASEDAMLAAFQLSRELRRLARKEPEFKPQYIALESLCQDYGFELLGMCRN
QSEVTAVLNDLGEDSETEPEAEGLGQAFEEGIPNLARLRLAVNYNQKQFVAHPICQQV
LSSIWCGNLAGWRGSTTIWRLFVASLIFLTMPFLCIGYWLAPKSQLGRLLKIPVLKFL
LHSASYLWFLIFLLGESLVMETQLSTFKGRSQSVWETSLHMIWVTGFLWFECKEVWIE
GLRSYLLDWWNFLDVVILSLYLASFALRLLLAGLAYMHCRDASDSTTCRCFTTAERSE
WRTEDPQFLAEVLFTVTSMLSFTRLAYILPAHESLGTLQISIGKMIDDMIRFMFILMI
ILTAFLCGLNNIYVPYQESEKLGNFNETFQFLFWTMFGMEEHTVVDMPQFLVPEFVGR
AMYGIFTIVMVIVLLNMLIAMITNSFQKIEDDADVEWKFARSKLYLSYFREGLTLPVP
FNILPSPKAAFYLVRRIFRFLCCGSSCCKAKKSDYPPIGTFTNPGARAGSAGEGERVS
YRLRVIKALVQRYIETARREFEETRRKDLGNRLTELTKTVSRLQSEVASVQKNLAAGG
APRPPDGASILSRYITRVRNSFQNLGPPTSDTPAELTMPGIVETEVSLGDGLDGTGEA
GAPAPGEPGSSSSAHVLVHREQEAEGSGDLLLEGDLETKGES"

BASE COUNT 769 a 1091 c 1014 g 817 t
ORIGIN

Query Match 67.8%; Score 15.6; DB 12; Length 3691;
Best Local Similarity 65.2%; Pred. No. 3.3e+02;
Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 gcngtngartggaayttygcnmg 23
| || ||:||||| ||:|| :|
Db 2883 GATGTGGAGTGGAAGTTTGCTCG 2905

Search completed: October 1, 1999, 15:34:44
Job time: 6187 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 1, 1999, 15:36:16 ; Search time 148.69 Seconds
(without alignments)
38.701 Million cell updates/sec

Title: US-09-026-400-5
Perfect score: 23
Sequence: 1 gcngtngartggaayttygcnmg 23

Scoring table: IDENTITY_NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		18.4	80.0	1660	1	V48147	Nicotianamine amin
2		18.4	80.0	23	1	V48149	Nicotianamine amin
3		17.2	74.8	1910	1	V48148	Nicotianamine amin
4		15.6	67.8	3258	1	V26030	Human transient re
5	c	15.6	67.8	29	1	V26038	Human transient re
6		15.4	67.0	11303	1	V52248	Streptococcus pneu
7		14.8	64.3	1193	1	V54602	Human uncoupling p
8		14.8	64.3	1220	1	V71710	Human uncoupling p
9		14.8	64.3	1033	1	V71711	Human uncoupling p
10		14.8	64.3	1231	1	V72690	Human uncoupling p
11		14.8	64.3	1132	1	V72691	Human uncoupling p
12		14.8	64.3	2340	1	V84254	Human uncoupling p
13	c	14.6	63.5	3094	1	N92579	Sequence of the 1.
14	c	14.6	63.5	2997	1	T93597	Eimeria tenella sp
15	c	14.6	63.5	3094	1	T93596	Eimeria tenella sp
16	c	14.2	61.7	4976	1	Q05754	Microspore-specifi
17	c	14.2	61.7	1938	1	Q05758	Clone representing
18	c	14.2	61.7	29879	1	Q46806	eryA region of S.
19		14.2	61.7	1054	1	Q99792	Plant SAR gene p11
20	c	14.2	61.7	4977	1	T48863	Brassica napus mic
21	c	14.2	61.7	1938	1	T48865	Brassica napus mic
22		14.2	61.7	3290	1	V26029	Human transient re
23	c	14.2	61.7	1344	1	T73169	Pseudomonas putida
24	c	14.2	61.7	1441	1	T73170	Pseudomonas putida
25	c	14.2	61.7	1417	1	T73171	Pseudomonas putida
26	c	14.2	61.7	1395	1	T73172	Pseudomonas putida
27	c	14.2	61.7	1374	1	T73173	Pseudomonas putida
28	c	14.2	61.7	1366	1	T73174	Pseudomonas putida
29	c	14.2	61.7	1412	1	T73175	Pseudomonas putida
30	c	14.2	61.7	1350	1	T73176	Pseudomonas putida
31	c	14.2	61.7	1349	1	T73177	Pseudomonas putida
32	c	14.2	61.7	1401	1	T73167	Pseudomonas putida
33	c	14.2	61.7	237326	1	V57903	Hereditary haemoch
34	c	14.2	61.7	235033	1	V57926	Hereditary haemoch
35		14.2	61.7	22080	1	X06751	Salmonella enteric
36		14	60.9	1266	1	T89425	Phospholipid scram
37		14	60.9	137507	1	V19941	KSHV long unique c
38		14	60.9	35100	1	V73804	KSHV LUR DNA (nucl
39		14	60.9	9626	1	X26819	Nucleotide sequenc
40	c	13.6	59.1	855	1	Q20622	C10-16 DNA fragmen
41	c	13.6	59.1	5270	1	Q25387	Rat thyrotropin re
42		13.6	59.1	1812	1	Q29419	Nicotiana tabacum
43	c	13.6	59.1	1918	1	Q31539	Lignin peroxidase
44	c	13.6	59.1	421	1	Q59416	Human brain Expres
45	c	13.6	59.1	491	1	Q62115	Superoxide dismuta

ALIGNMENTS

RESULT 1
V48147
ID V48147 standard; cDNA to mRNA; 1660 BP.
AC V48147;
DT 27-OCT-1998 (first entry)
DE Nicotianamine aminotransferase 49564.15 molecular weight protein, gene.
KW ds; nicotianamine aminotransferase; plant; iron absorption;
KW iron deficiency chlorosis.
OS Gramineae sp.
FH Key Location/Qualifiers
FT CDS 62. .1447
FT /*tag= a
FT /product= "Nicotianamine aminotransferase"
PN EP-860499-A2.
PD 26-AUG-1998.
PF 19-FEB-1998; 102891.
PR 21-FEB-1997; JP-037499.
PA (SUMO) SUMITOMO CHEM CO LTD.
PI Mori S, Nakanishi H, Takahashi M;
DR WPI; 98-439341/38.
DR P-PSDB; W61642.
PT New nicotianamine aminotransferase protein and DNA - useful for
PT enhancing iron absorption of plant cells
PS Claim 4; Page 12-13; 17pp; English.
CC The nicotianamine aminotransferase can be used in a plasmid to transform
CC plant cells to produce cells with enhanced iron absorption, and it is
CC implied [though not stated] that plants with improved resistance to iron
CC deficiency chlorosis in calcareous soils can be regenerated from the
CC transformed cells. The gene fragment can be used to detect, amplify
CC and/or isolate nicotianamine aminotransferase genes.
SQ Sequence 1660 BP; 423 A; 442 C; 430 G; 365 T;

Query Match 80.0%; Score 18.4; DB 1; Length 1660;
Best Local Similarity 69.6%; Pred. No. 0.89;
Matches 16; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 gcngtngartggaayttygcnmg 23
          || || ||:|||||:||:|| :|
Db    176 GCGGTGGAGTGGAATTTCGCCCG 198
```

RESULT 2
V48149
ID V48149 standard; DNA ; 23 BP.
AC V48149;
DT 27-OCT-1998 (first entry)
DE Nicotianamine aminotransferase primer 1.
KW ss; nicotianamine aminotransferase; plant; iron absorption; PCR;
KW iron deficiency chlorosis; primer; amplification.
OS Synthetic.
OS Gramineae sp.
FH Key Location/Qualifiers
FT modified_base 3
FT /*tag= a
FT /mod_base= I
FT modified_base 6
FT /*tag= b
FT /mod_base= I
FT modified_base 21
FT /*tag= c
FT /mod_base= I
PN EP-860499-A2.
PD 26-AUG-1998.
PF 19-FEB-1998; 102891.
PR 21-FEB-1997; JP-037499.

PA (SUMO) SUMITOMO CHEM CO LTD.
PI Mori S, Nakanishi H, Takahashi M;
DR WPI; 98-439341/38.
PT New nicotianamine aminotransferase protein and DNA - useful for
PT enhancing iron absorption of plant cells
PS Claim 16; Page 7; 17pp; English.
CC The primers V48149 and V48150 were used in the cloning of nicotianamine
CC aminotransferase can be used in a plasmid to transform plant cells to
CC produce cells with enhanced iron absorption, and it is implied [though
CC not stated] that plants with improved resistance to iron deficiency
CC chlorosis in calcareous soils can be regenerated from the transformed
CC cells. The gene fragment can be used to detect, amplify and/or isolate
CC nicotianamine aminotransferase genes.
SQ Sequence 23 BP; 3 A; 2 C; 7 G; 4 T;

Query Match 80.0%; Score 18.4; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gcngtngartggaayttygcnmg 23
||||||||||||||||||||||
Db 1 GCNGTNGARTGGAAYTTYGCNMG 23

RESULT 3
V48148
ID V48148 standard; cDNA to mRNA; 1910 BP.
AC V48148;
DT 27-OCT-1998 (first entry)
DE Nicotianamine aminotransferase 58148.62 molecular weight protein, gene.
KW ds; nicotianamine aminotransferase; plant; iron absorption;
KW iron deficiency chlorosis.
OS Gramineae sp.
FH Key Location/Qualifiers
FT CDS 76. .1731
FT /*tag= a
FT /product= "Nicotianamine transferase"
PN EP-860499-A2.
PD 26-AUG-1998.
PF 19-FEB-1998; 102891.
PR 21-FEB-1997; JP-037499.
PA (SUMO) SUMITOMO CHEM CO LTD.
PI Mori S, Nakanishi H, Takahashi M;
DR WPI; 98-439341/38.
DR P-PSDB; W61643.
PT New nicotianamine aminotransferase protein and DNA - useful for
PT enhancing iron absorption of plant cells
PS Claim 4; Page 14-15; 17pp; English.
CC The nicotianamine aminotransferase can be used in a plasmid to transform
CC plant cells to produce cells with enhanced iron absorption, and it is
CC implied [though not stated] that plants with improved resistance to iron
CC deficiency chlorosis in calcareous soils can be regenerated from the
CC transformed cells. The gene fragment can be used to detect, amplify
CC and/or isolate nicotianamine aminotransferase genes.
SQ Sequence 1910 BP; 462 A; 534 C; 546 G; 368 T;

Query Match 74.8%; Score 17.2; DB 1; Length 1910;
Best Local Similarity 69.6%; Pred. No. 3.9;
Matches 16; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gcngtngartggaayttygcnmg 23
|| || ||:|||||:||:|| |
Db 460 GCGGTGGAGTGGAATTTCGCGGG 482

RESULT 4
V26030
ID V26030 standard; cDNA; 3258 BP.
AC V26030;
DT 28-AUG-1998 (first entry)
DE Human transient receptor potential protein Htrp3 cDNA.
KW Htrp3; transient receptor potential; trp protein; human;
KW capacitative calcium ion entry; CCE; asthma; hypertension;
KW diabetes; osteoporosis; osteogenesis; thrombosis; immunodeficiency;
KW gene therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 291. .3080
FT /*tag= a
PN WO9808979-A1.
PD 05-MAR-1998.
PF 29-AUG-1997; U15247.
PR 15-OCT-1996; US-729955.
PR 30-AUG-1996; US-025111.
PA (REGC) UNIV CALIFORNIA.
PI Birnbaumer L, Zhu X;
DR WPI; 98-230269/20.
DR P-PSDB; W55961.
PT Controlling capacitative calcium ion entry into mammalian cells - by
PT changing activity of transient receptor potential proteins, e.g. for
PT treating asthma, hypertension etc.
PS Claim 12; Page 31-35; 60pp; English.
CC This cDNA clone codes for a human transient receptor potential
CC (trp) protein (see W55961), designated Htrp3, that is an essential
CC part of the capacitative calcium ion entry (CCE) mechanism in human
CC cells. The cDNA was isolated by subjecting human embryo cell line
CC 293 cDNA to RACE-PCR using primers (see V26035-38) based on the
CC sequence of GeneBank expressed sequence tag EST R34716. Htrp1 (see
CC V26029) cDNA has also been isolated. CCE into a mammalian cell
CC expressing a trp protein required for CCE is controlled in a
CC claimed method by treating the cell with an agent that increases or
CC decreases the amount of biologically active trp protein from its
CC normal level. Agents that inhibit CCE are potentially useful for
CC treating asthma, hypertension and osteoporosis, also for
CC antithrombotic therapy, while those that stimulate CCE are used to
CC treat type II diabetes and to induce bone formation. Primary
CC immunodeficiency, if associated with trp gene mutations, may be
CC treated by gene therapy.
SQ Sequence 3258 BP; 911 A; 720 C; 787 G; 840 T;

Query Match 67.8%; Score 15.6; DB 1; Length 3258;
Best Local Similarity 65.2%; Pred. No. 29;
Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 gcngtngartggaayttygcnmg 23
| || ||:|||| ||:|| :|
Db 2502 GACGTGGAGTGGAAGTTTGCAAG 2524

RESULT 5
V26038/c
ID V26038 standard; DNA; 29 BP.
AC V26038;
DT 28-AUG-1998 (first entry)
DE Human transient receptor potential protein Htrp3 primer A2.
KW Htrp3; transient receptor potential; trp protein; human;
KW capacitative calcium ion entry; CCE; asthma; hypertension;
KW diabetes; osteoporosis; osteogenesis; thrombosis; immunodeficiency;
KW gene therapy; RACE; PCR; primer; ss.
OS Synthetic.
OS Homo sapiens.
PN WO9808979-A1.
PD 05-MAR-1998.
PF 29-AUG-1997; U15247.
PR 15-OCT-1996; US-729955.
PR 30-AUG-1996; US-025111.
PA (REGC) UNIV CALIFORNIA.
PI Birnbaumer L, Zhu X;
DR WPI; 98-230269/20.
PT Controlling capacitative calcium ion entry into mammalian cells - by
PT changing activity of transient receptor potential proteins, e.g. for
PT treating asthma, hypertension etc.
PS Disclosure; Page 9; 60pp; English.

```
CC   PCR primers S1, S2, A1 and A2 (see V26035-38) are based on the
CC   expressed sequence tag EST R34716 from the GenBank dbEST database.
CC   They were used in the isolation and identification of Htrp3 cDNA.
CC   Primary RACE-PCR amplifications of human embryonic kidney (HEK) 293
CC   cell cDNA was performed using adaptor primer AP1 in combination
CC   with primer S1 for 3' amplification or A1 for 5' amplification.
CC   Nested-PCR amplifications were performed using internal primers
CC   AP2 plus S2 for 3'RACE or AP2 plus A2 for 5'RACE.  Positive
CC   clones were identified using oligonucleotides A1 and S2 for the
CC   3' and 5' RACE respectively.  Htrp3 cDNA (see V26030) was obtained.
CC   It codes for a transient receptor potential (trp) protein (see
CC   W55961) that is essential for calcium ion entry into cells.  The
CC   invention relates to methods for treating cells with agents that
CC   raise or lower the amount of trp protein and which thereby control
CC   capacitative calcium ion entry into the cell, and for using trp
CC   proteins as targets for identifying such agents.
SQ   Sequence   29 BP;    8 A;    10 C;    3 G;    8 T;


  Query Match             67.8%;  Score 15.6;  DB 1;  Length 29;
  Best Local Similarity   65.2%;  Pred. No. 14;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 gcngtngartggaayttygcnmg 23
           |  || ||:||||| ||:|| :|
Db      23 GATGTAGAATGGAAGTTTGCTCG 1


RESULT   6
V52248
ID   V52248 standard; DNA; 11303 BP.
AC   V52248;
DT   23-OCT-1998  (first entry)
DE   Streptococcus pneumoniae genome fragment SEQ ID NO:115.
KW   Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW   computer readable medium; vaccine; pharmaceutical composition; ds.
OS   Streptococcus pneumoniae.
PN   WO9818931-A2.
PD   07-MAY-1998.
PF   30-OCT-1997; U19588.
PR   31-OCT-1996; US-029960.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M,
PI   Kunsch CA, Rosen CA;
DR   WPI; 98-272225/24.
PT   Computer-readable medium with recorded Streptococcus pneumoniae
PT   polynucleotide sequences - useful in diagnostic kits and assays, and
PT   pharmaceutical compositions and vaccines for Streptococcus
PT   pneumoniae
PS   Claim 1; Page 831-838; 1409pp; English.
CC   The present invention describes a computer readable medium which has
CC   the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
CC   on it, or a representative fragment or a sequence at least 95% identical
CC   to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC   to 391 (V52134 to V52524) are genomic fragments from Streptococcus
CC   pneumoniae. The present invention also describes an isolated nucleic acid
CC   molecule encoding a homologue of any of the fragments of the S.pneumoniae
CC   genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC   by a process comprising: (a) screening a genomic DNA library using as a
CC   probe a target sequence defined by any of the sequences in SEQ ID NO:1
CC   to 391, identifying members of the library which contain sequences
CC   that hybridise to the target sequence and isolating the nucleic acid
CC   molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
CC   from an organism, amplifying nucleic acid molecules whose nucleotide
CC   sequence is homologous to amplification primers derived from the
CC   fragment of the S. pneumoniae genome to prime the amplification and
CC   isolating the amplified sequences. The computer readable medium can be
CC   used in a computer-based system for identifying fragments of the
CC   S. pneumoniae genome of commercial importance, or expression modulating
CC   fragments of the S. pneumoniae genome. Products from the present
CC   invention can be used in diagnosis kits and assays, and pharmaceutical
CC   compositions and vaccines for S. pneumoniae.
SQ   Sequence   11303 BP;    3733 A;    2266 C;    1558 G;    3746 T;


  Query Match             67.0%;  Score 15.4;  DB 1;  Length 11303;
  Best Local Similarity   68.4%;  Pred. No. 45;
  Matches   13;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 gtngartggaayttygcnm 22
           || ||:|||||:|||:|| :
Db     189 GTCGAATGGAATTTCGCTA 207


RESULT   7
V54602
ID   V54602 standard; cDNA; 1193 BP.
AC   V54602;
DT   07-DEC-1998  (first entry)
DE   Human uncoupling protein HNFCW60 cDNA.
KW   HNFCW60; uncoupling protein; human; body weight disorder; obesity;
KW   diabetes; hyperlipidaemia; diagnosis; therapy; vaccine; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             199. .1137
FT                   /*tag= a
PN   WO9839432-A1.
PD   11-SEP-1998.
PF   02-MAR-1998; G00633.
PR   16-JUL-1997; EP-305305.
PR   05-MAR-1997; GB-004551.
PR   18-MAR-1997; GB-005614.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
PI   Beeley LJ, Godden RJ, Paine K;
DR   WPI; 98-495841/42.
DR   P-PSDB; W68197.
PT   New isolated human uncoupling polypeptide(s) - used to develop
PT   products for the diagnosis, prevention and treatment of body weight
PT   disorders, obesity and diabetes
PS   Claim 3; Page 24-25; 41pp; English.
CC   This isolated polynucleotide (PN) codes for novel human uncoupling
CC   protein HNFCW60 (see W68197).  HNFCW60 polynucleotides (see also
CC   V54603-04) can be obtained from a cDNA library derived from mRNA in
CC   cells of human brain frontal cortex, rhabdomyosarcoma, foetal heart
CC   and skeletal muscle using expressed sequence tag analysis.  The
CC   invention relates to HNFCW60 polypeptides and recombinant materials
CC   and methods for their production.  It also relates to methods for
CC   using such HNFCW60 polypeptides and polynucleotides.  Such uses
CC   include the treatment of obesity, diabetes, hyperlipidaemia and
CC   body weight disorders.  The invention also provides methods to
CC   identify agonists and antagonists, and methods for treating
CC   conditions associated with HNFCW60 imbalance using the identified
CC   compounds.  In addition, diagnostic assays for detecting diseases
CC   associated with inappropriate HNFCW60 activity or levels are
CC   provided.
SQ   Sequence   1193 BP;    252 A;    370 C;    343 G;    228 T;


  Query Match             64.3%;  Score 14.8;  DB 1;  Length 1193;
  Best Local Similarity   61.9%;  Pred. No. 66;
  Matches   13;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 cngtngartggaayttygcnm 22
           | || ||: ||||:||:|| :
Db     730 CTGTGGAAAGGAACTTTGCCC 750


RESULT   8
V71710
ID   V71710 standard; DNA; 1220 BP.
AC   V71710;
DT   09-FEB-1999  (first entry)
DE   Human uncoupling protein 3 (UCP3) encoding DNA.
KW   Uncoupling protein 3; UCP3; thermogenesis; mammal; enhancer; drug;
KW   protein catabolism; anti-obesity; inhibitor; muscle wasting; infection;
KW   HIV; cancer; tumour cachexia; muscle disease; muscular dystrophy;
```

```
KW   non-insulin dependent diabetes mellitus; diagnosis; human; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             184. .1122
FT                   /*tag=  a
FT                   /product= "human UCP3"
FT   CDS             1. .1220
FT                   /*tag=  b
FT                   /note= "reading frame 1; contains internal stop codons
FT                          which are not indicated in the corresponding
FT                          protein (W81588)"
FT   CDS             2. .1219
FT                   /*tag=  c
FT                   /note= "reading frame 2; contains internal stop codons
FT                          which are not indicated in the corresponding
FT                          protein (W81589)"
FT   CDS             3. .1220
FT                   /*tag=  d
FT                   /note= "reading frame 3; contains internal stop codons
FT                          which are not indicated in the corresponding
FT                          protein (W81590)"
PN   WO9845438-A1.
PD   15-OCT-1998.
PF   08-APR-1998; U06959.
PR   15-JUL-1997; US-892745.
PR   09-APR-1997; US-043447.
PR   12-MAY-1997; US-046254.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PI   Flier JS, Lowell BB;
DR   WPI; 98-594483/50.
DR   P-PSDB; W81587, W81588, W81589, W81590.
PT   New isolated uncoupling protein, UCP-3 - used to develop products
PT   for modulating thermogenesis in tissues, e.g. for treating obesity
PT   or muscle wasting caused by infection or cancer
PS   Claim 3; Fig 1A-C; 98pp; English.
CC   This DNA encodes a human uncoupling protein 3 (UCP3). A host cell
CC   transformed with a construct comprising the UCP3 nucleic acid can be used
CC   for the recombinant production of the protein. The UCP3 is involved in
CC   the regulation of thermogenesis in mammals. The nucleic acids (V71710 to
CC   V71712) can be used for identifying compounds which alter UCP3 activity.
CC   Enhancers of UCP3 can be used for enhancing protein catabolism in a
CC   mammal and can be used as anti-obesity drugs. Inhibitors of UCP3 can be
CC   used for inhibiting protein catabolism in a mammal such as inhibiting
CC   muscle wasting. They can be used for curtailing muscle wasting due to
CC   infection (e.g. HIV), cancer, tumour cachexia, muscle diseases (e.g.
CC   muscular dystrophy) or as a possible treatment for non-insulin dependent
CC   diabetes mellitus. The products can also be used for detection and
CC   diagnosis.
SQ   Sequence   1220 BP;    267 A;    376 C;    345 G;    232 T;

  Query Match             64.3%;  Score 14.8;  DB 1;  Length 1220;
  Best Local Similarity   61.9%;  Pred. No. 66;
  Matches   13;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 cngtngartggaayttygcnm 22
           | || ||: ||||:||:|| :
Db     715 CTGTGGAAAGGAACTTTGCCC 735


RESULT   9
V71711
ID   V71711 standard; DNA; 1033 BP.
AC   V71711;
DT   09-FEB-1999  (first entry)
DE   Human uncoupling protein 3 short form (UCP3sh) encoding DNA.
KW   Uncoupling protein 3; UCP3; thermogenesis; mammal; enhancer; drug;
KW   protein catabolism; anti-obesity; inhibitor; muscle wasting; infection;
KW   HIV; cancer; tumour cachexia; muscle disease; muscular dystrophy;
KW   non-insulin dependent diabetes mellitus; diagnosis; human; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             181. .1008
FT                   /*tag=  a
FT                   /product= "human UCP3"
FT   CDS             1. .1032
FT                   /*tag=  b
FT                   /note= "reading frame 1; contains internal stop codons
FT                          which are not indicated in the corresponding
FT                          protein (W81592)"
FT   CDS             2. .1033
FT                   /*tag=  c
FT                   /note= "reading frame 2; contains internal stop codons
FT                          which are not indicated in the corresponding
FT                          protein (W81593)"
FT   CDS             3. .1033
FT                   /*tag=  d
FT                   /note= "reading frame 3; contains internal stop codons
FT                          which are not indicated in the corresponding
FT                          protein (W81594)"
PN   WO9845438-A1.
PD   15-OCT-1998.
PF   08-APR-1998; U06959.
PR   15-JUL-1997; US-892745.
PR   09-APR-1997; US-043447.
PR   12-MAY-1997; US-046254.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PI   Flier JS, Lowell BB;
DR   WPI; 98-594483/50.
DR   P-PSDB; W81591, W81592, W81593, W81594.
PT   New isolated uncoupling protein, UCP-3 - used to develop products
PT   for modulating thermogenesis in tissues, e.g. for treating obesity
PT   or muscle wasting caused by infection or cancer
PS   Claim 3; Fig 2A-B; 98pp; English.
CC   This DNA encodes a human uncoupling protein 3 short form (UCP3sh).
CC   The invention provides human and mouse UCP3 genes (V71710 and V71712)
CC   encoding UCP3 proteins (W81587 and W81595) respectively. A host cell
CC   transformed with a construct comprising the UCP3 nucleic acid can be used
CC   for the recombinant production of the protein. The UCP3 is involved in
CC   the regulation of thermogenesis in mammals. The nucleic acids (V71710 to
CC   V71712) can be used for identifying compounds which alter UCP3 activity.
CC   Enhancers of UCP3 can be used for enhancing protein catabolism in a
CC   mammal and can be used as anti-obesity drugs. Inhibitors of UCP3 can be
CC   used for inhibiting protein catabolism in a mammal such as inhibiting
CC   muscle wasting. They can be used for curtailing muscle wasting due to
CC   infection (e.g. HIV), cancer, tumour cachexia, muscle diseases (e.g.
CC   muscular dystrophy) or as a possible treatment for non-insulin dependent
CC   diabetes mellitus. The products can also be used for detection and
CC   diagnosis.
SQ   Sequence   1033 BP;    204 A;    340 C;    297 G;    192 T;

  Query Match             64.3%;  Score 14.8;  DB 1;  Length 1033;
  Best Local Similarity   61.9%;  Pred. No. 64;
  Matches   13;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 cngtngartggaayttygcnm 22
           | || ||: ||||:||:|| :
Db     712 CTGTGGAAAGGAACTTTGCCC 732


RESULT  10
V72690
ID   V72690 standard; cDNA; 1231 BP.
AC   V72690;
DT   22-FEB-1999  (first entry)
DE   Human uncoupling protein UCP3L encoding cDNA.
KW   Human; uncoupling protein; UCP3L; UCP3S; oxidative phosphorylation;
KW   adenosine triphosphate; mitochondria; skeletal muscle; gene therapy;
KW   thermogenesis; heart; obesity; cachexia; type II diabetes; hypertension;
KW   insulin sensitivity; neuromuscular disease; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             154. .1092
FT                   /*tag=  a
FT                   /product= "UCP3L"
```

```
PN   WO9850542-A1.
PD   12-NOV-1998.
PF   05-MAY-1998; E02645.
PR   07-MAY-1997; CH-001072.
PA   (NOVS ) NOVARTIS AG.
PA   (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PI   Boss O, Giacobino J, Muzzin P;
DR   WPI; 98-610382/51.
DR   P-PSDB; W83379.
PT   New human skeletal muscle uncoupling proteins UCP3L and UCP3S -
PT   useful for controlling thermogenesis in human skeletal muscle and
PT   heart, e.g. for treating obesity and cachexia
PS   Claim 1; Page 14-15; 26pp; English.
CC   The present sequence encodes human uncoupling protein UCP3L. UCP3
CC   uncouples oxidative phosphorylation and synthesis of adenosine
CC   triphosphate in the mitochondria of skeletal muscle. The coding
CC   sequences for UCP3L and UCP3S are useful for gene therapy of
CC   dysfunctions of thermogenesis in human skeletal muscle and heart which
CC   result from a lack of UCP3 and which can induce disorders such as
CC   obesity or cachexia. Antisense oligonucleotides to UCP3L and UCP3S can
CC   be used for correcting an excess of UCP3. Modification of endogenous
CC   UCP3 activity (using activators or inhibitors of UCP3) is used to induce
CC   bodyweight loss (loss of adipose mass and maintenance of the lean mass)
CC   in all types of obesity by promoting the dissipation of energy; for
CC   preventing an excessive weight regain following restrictive food diet or
CC   after ceasing a physical training programme; for preventing and treating
CC   type II diabetes by improving sensitivity to insulin; for preventing
CC   hypertension; for increasing muscle mass in states of cachexia; for
CC   treatment of insufficiencies or disturbances of cardiac rhythm due to a
CC   dysfunction of UCP3; and for the treatment of neuromuscular diseases due
CC   to a dysfunction of UCP3. The uncoupling proteins can also be used to
CC   raise antibodies, e.g. for diagnosis. Knowledge of the UCP3 genes allows
CC   generation of transgenic animals, e.g. for screening substances which
CC   modify UCP3 expression or activity or for investigating the biological
CC   role of UCP3.
SQ   Sequence   1231 BP;    271 A;    371 C;    345 G;    244 T;


  Query Match             64.3%;  Score 14.8;  DB 1;  Length 1231;
  Best Local Similarity   61.9%;  Pred. No. 66;
  Matches   13;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnm 22
          | || ||: ||||:||:|| :
Db    685 CTGTGGAAAGGAACTTTGCCC 705


RESULT  11
V72691
ID   V72691 standard; cDNA; 1132 BP.
AC   V72691;
DT   22-FEB-1999  (first entry)
DE   Human uncoupling protein UCP3S encoding cDNA.
KW   Human; uncoupling protein; UCP3L; UCP3S; oxidative phosphorylation;
KW   adenosine triphosphate; mitochondria; skeletal muscle; gene therapy;
KW   thermogenesis; heart; obesity; cachexia; type II diabetes; hypertension;
KW   insulin sensitivity; neuromuscular disease; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             154. .991
FT                   /*tag= a
FT                   /product= "UCP3S"
PN   WO9850542-A1.
PD   12-NOV-1998.
PF   05-MAY-1998; E02645.
PR   07-MAY-1997; CH-001072.
PA   (NOVS ) NOVARTIS AG.
PA   (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PI   Boss O, Giacobino J, Muzzin P;
DR   WPI; 98-610382/51.
DR   P-PSDB; W83380.
PT   New human skeletal muscle uncoupling proteins UCP3L and UCP3S -
PT   useful for controlling thermogenesis in human skeletal muscle and
PT   heart, e.g. for treating obesity and cachexia
PS   Claim 3; Page 17-18; 26pp; English.
CC   The present sequence encodes human uncoupling protein UCP3S. UCP3
CC   uncouples oxidative phosphorylation and synthesis of adenosine
CC   triphosphate in the mitochondria of skeletal muscle. The coding
CC   sequences for UCP3L and UCP3S are useful for gene therapy of
CC   dysfunctions of thermogenesis in human skeletal muscle and heart which
CC   result from a lack of UCP3 and which can induce disorders such as
CC   obesity or cachexia. Antisense oligonucleotides to UCP3L and UCP3S can
CC   be used for correcting an excess of UCP3. Modification of endogenous
CC   UCP3 activity (using activators or inhibitors of UCP3) is used to induce
CC   bodyweight loss (loss of adipose mass and maintenance of the lean mass)
CC   in all types of obesity by promoting the dissipation of energy; for
CC   preventing an excessive weight regain following restrictive food diet or
CC   after ceasing a physical training programme; for preventing and treating
CC   type II diabetes by improving sensitivity to insulin; for preventing
CC   hypertension; for increasing muscle mass in states of cachexia; for
CC   treatment of insufficiencies or disturbances of cardiac rhythm due to a
CC   dysfunction of UCP3; and for the treatment of neuromuscular diseases due
CC   to a dysfunction of UCP3. The uncoupling proteins can also be used to
CC   raise antibodies, e.g. for diagnosis. Knowledge of the UCP3 genes allows
CC   generation of transgenic animals, e.g. for screening substances which
CC   modify UCP3 expression or activity or for investigating the biological
CC   role of UCP3.
SQ   Sequence   1132 BP;    245 A;    351 C;    321 G;    215 T;


  Query Match             64.3%;  Score 14.8;  DB 1;  Length 1132;
  Best Local Similarity   61.9%;  Pred. No. 65;
  Matches   13;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnm 22
          | || ||: ||||:||:|| :
Db    685 CTGTGGAAAGGAACTTTGCCC 705


RESULT  12
V84254
ID   V84254 standard; cDNA; 2340 BP.
AC   V84254;
DT   12-APR-1999  (first entry)
DE   Human uncoupling protein 3 (UCP3) cDNA.
KW   Uncoupling protein 3; UCP3; human; obesity; diabetes;
KW   hyperinsulinaemia; hypermetabolism; gene therapy; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             344. .1282
FT                   /*tag= a
FT                   /note= "this region is specifically claimed in
FT                         Claim 2"
PN   WO9852958-A1.
PD   26-NOV-1998.
PF   19-MAY-1998; U10215.
PR   09-DEC-1997; US-069141.
PR   20-MAY-1997; US-047179.
PR   08-AUG-1997; GB-016886.
PA   (MERI ) MERCK & CO INC.
PI   Chen F, Liu Q;
DR   WPI; 99-059737/05.
DR   P-PSDB; W88279.
PT   New isolated human uncoupling protein 3 - used to identify compounds
PT   which can modulate energy expenditure and body weight regulation,
PT   e.g. in the treatment of obesity or diabetes.
PS   Claim 1; Page 57-58; 87pp; English.
CC   This cDNA clone encodes human uncoupling protein 3 (UCP3, see
CC   W88279), a novel protein that is involved in energy expenditure and
CC   body weight regulation and whose expression is mostly limited to
CC   skeletal muscle.  The clone was isolated from a foetal brain cDNA
CC   library using primers (see V84259-64) based on isolated EST clones
CC   (see V84255-58) and pBluescript vector sequences.  The invention
CC   additionally provides related recombinant expression vectors,
CC   recombinant host cells and purified forms of the UCP3 protein.
CC   The UCP3 polypeptides and transformed recombinant cell lines can
```

```
CC   be used for identifying modulators of UCP3 activity.  Such
CC   modulators can be used for treating diseases such as obesity and
CC   diabetes, by manipulating the interrelated process of balancing
CC   food intake, energy expenditure and glucose metabolism within the
CC   patient.  They can also be used to treat hyperactive conditions of
CC   energy expenditure which originate in the mitochondria of skeletal
CC   muscle.  UCP3 nucleic acids are useful in gene therapy of obesity
CC   and obesity-related indications, including diabetes, and of
CC   mitochondrial-associated hypermetabolism.
SQ   Sequence   2340 BP;    606 A;    638 C;    633 G;    463 T;

  Query Match             64.3%;  Score 14.8;  DB 1;  Length 2340;
  Best Local Similarity   61.9%;  Pred. No. 73;
  Matches   13;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 cngtngartggaayttygcnm 22
           | || ||: ||||:||:|| :
Db     875 CTGTGGAAAGGAACTTTGCCC 895


RESULT  13
N92579/c
ID   N92579 standard; DNA; 3094 BP.
AC   N92579;
DT   15-MAR-1992  (first entry)
DE   Sequence of the 1.7kb cDNA molecule encoding antibodies 7D1,7D4
DE   and 20C6.
KW   Vaccine; coccidiosis; poultry; bivalent vaccine; ss.
OS   Eimeria.
FH   Key             Location/Qualifiers
FT   cds             9. .3002
FT                   /*tag=  a
PN   EP-344808-A.
PD   06-DEC-1989.
PF   02-JUN-1989; 110056.
PR   03-JUN-1988; US-202721.
PA   (HOFF ) HOFFMANN-LA ROCHE AG.
PI   Altenburger W, Binger MH, Chizzonite RA, Kramer RA,
PI   Lomedico PT, McAndrew SJ;
DR   WPI; 89-358220/49.
DR   P-PSDB; P93706.
PT   DNA sequences encoding Eimeria surface antigens - used in
PT   pox:virus vectors as a vaccine to protect chicks against
PT   coccidiosis
PS   Claim 10; Fig 20A-D; 78pp; English.
CC   The inventors claim a new protein which comprises one or more
CC   immunoreactive and/or antigenic determinants of an Eimeria surface
CC   antigen of mol. wt. 28,37,120 or more than 200 kD which specifically
CC   binds to one or more monoclonal antibody (MaB) from ATCC HB 9707-9712
CC   (see P93703-6). Also new are DNA encoding the protein (see N92576-9),
CC   and a vaccine comprising one or more proteins. Vaccine utility can
CC   be enhanced by inserting additional genes into the carrier virus
CC   (see P91652).
SQ   Sequence   3094 BP;    617 A;    834 C;    846 G;    796 T;

  Query Match             63.5%;  Score 14.6;  DB 1;  Length 3094;
  Best Local Similarity   63.6%;  Pred. No. 97;
  Matches   14;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 gcngtngartggaayttygcnm 22
           || |  ||:|||||  ||:|| :
Db     262 GCAGGGGAGTGGAAGTTCGCGA 241


RESULT  14
T93597/c
ID   T93597 standard; cDNA; 2997 BP.
AC   T93597;
DT   21-MAY-1998  (first entry)
DE   Eimeria tenella sporozoite 45 kDa surface antigen cDNA.
KW   Coccidiosis; vaccine; poultry; protozoan; parasite; antigen;
KW   sporozoite; ss.
OS   Eimeria tenella.
FH   Key             Location/Qualifiers
FT   CDS             7. .2997
FT                   /*tag=  a
PN   US5661015-A.
PD   26-AUG-1997.
PF   03-JUN-1988; 202721.
PR   20-DEC-1991; US-812349.
PR   03-JUN-1988; US-202721.
PA   (HOFF ) HOFFMANN LA ROCHE INC.
PI   Altenburger W, Binger M, Chizzonite RA, Kramer RA,
PI   Lomedico PT, McAndrew SJ;
DR   WPI; 97-434379/40.
DR   P-PSDB; W33624.
PT   New DNA from Eimeria tenella and related immunogenic polypeptides -
PT   useful in vaccines to protect poultry against coccidiosis
PS   Claim 2; Fig 33A-B; 72pp; English.
CC   This cDNA sequence comprises the coding region for a 45 kDa protein
CC   (see W33624) that is recognised by monoclonal antibody 7B2
CC   (ATCC HB 9712).  This antibody also specifically reacts with an
CC   Eimeria tenella 200 kDa surface antigen that is present in the
CC   sporozoite developmental stage.  The clone (see also T93596) was
CC   obtained from a cDNA library by immunological screening with
CC   monoclonal antibodies raised against Eimeria antigens.  The
CC   invention provides DNA sequences (see T93593-98) coding for Eimeria
CC   surface antigens (see W31582-84 and W33621-26), recombinant vectors
CC   containing such DNA sequences, transformed microorganisms
CC   containing such vectors, and methods for producing the antigens
CC   using the transformed microorganisms.  Methods are also provided
CC   for protecting poultry against coccidiosis using the Eimeria
CC   surface antigens.  The surface antigens are administered either as
CC   purified proteins or in the form of DNA encoding the proteins in
CC   a viral vector such as a vaccinia virus.  The vaccines may produce
CC   antibodies that are cross-reactive with other Eimeria species.
SQ   Sequence   2997 BP;    583 A;    815 C;    830 G;    769 T;

  Query Match             63.5%;  Score 14.6;  DB 1;  Length 2997;
  Best Local Similarity   63.6%;  Pred. No. 97;
  Matches   14;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 gcngtngartggaayttygcnm 22
           || |  ||:|||||  ||:|| :
Db     254 GCAGGGGAGTGGAAGTTCGCGA 233


RESULT  15
T93596/c
ID   T93596 standard; cDNA; 3094 BP.
AC   T93596;
DT   21-MAY-1998  (first entry)
DE   Eimeria tenella sporozoite 45 kDa surface antigen cDNA.
KW   Coccidiosis; vaccine; poultry; protozoan; parasite; antigen;
KW   sporozoite; ss.
OS   Eimeria tenella.
FH   Key             Location/Qualifiers
FT   CDS             9. .3005
FT                   /*tag=  a
FT                   /transl_except= (pos:600. .602, aa:Tyr)
FT                   /transl_except= (pos:603. .605, aa:Ser)
FT                   /transl_except= (pos:606. .608, aa:Gly)
FT                   /transl_except= (pos:612. .614, aa:Ala)
FT                   /transl_except= (pos:615. .617, aa:Leu)
FT                   /transl_except= (pos:618. .620, aa:Ala)
FT   misc_difference 2001
FT                   /*tag=  b
FT                   /note= "base 2001 is given as '8' in Fig20
FT                         of the specification"
PN   US5661015-A.
PD   26-AUG-1997.
PF   03-JUN-1988; 202721.
PR   20-DEC-1991; US-812349.
```

PR 03-JUN-1988; US-202721.
PA (HOFF) HOFFMANN LA ROCHE INC.
PI Altenburger W, Binger M, Chizzonite RA, Kramer RA,
PI Lomedico PT, McAndrew SJ;
DR WPI; 97-434379/40.
DR P-PSDB; W33621.
PT New DNA from Eimeria tenella and related immunogenic polypeptides -
PT useful in vaccines to protect poultry against coccidiosis
PS Claim 1; Fig 20A-D; 72pp; English.
CC This cDNA clone includes a coding region for a 45 kDa protein (see
CC W33621) that is recognised by monoclonal antibody 7B2 (ATCC HB 9712).
CC This antibody also specifically reacts with an Eimeria tenella 200
CC kDa surface antigen that is present in the sporozoite developmental
CC stage. The clone was obtained from a cDNA library by immunological
CC screening with monoclonal antibodies raised against Eimeria
CC antigens. The first and last 7 nucleotides of the sequence are
CC derived from linker sequences used in the cloning procedure. The
CC invention provides DNA sequences (see T93593-98) coding for Eimeria
CC surface antigens (see W31582-84 and W33621-26), recombinant vectors
CC containing such DNA sequences, transformed microorganisms
CC containing such vectors, and methods for producing the antigens
CC using the transformed microorganisms. Methods are also provided
CC for protecting poultry against coccidiosis using the Eimeria
CC surface antigens. The surface antigens are administered either as
CC purified proteins or in the form of DNA encoding the proteins in
CC a viral vector such as a vaccinia virus. The vaccines may produce
CC antibodies that are cross-reactive with other Eimeria species.
SQ Sequence 3094 BP; 619 A; 834 C; 844 G; 796 T;

Query Match 63.5%; Score 14.6; DB 1; Length 3094;
Best Local Similarity 63.6%; Pred. No. 97;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 gcngtngartggaayttygcnm 22
           || |  ||:||||| ||:|| :
Db     262 GCAGGGGAGTGGAAGTTCGCGA 241
```

Search completed: October 1, 1999, 15:36:18
Job time: 6116 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 1, 1999, 15:03:35 ; Search time 1096.08 Seconds
(without alignments)
41.391 Million cell updates/sec

Title: US-09-026-400-5
Perfect score: 23
Sequence: 1 gcngtngartggaayttygcnmg 23

Scoring table: IDENTITY_NUC

Searched: 2546578 seqs, 986266752 residues

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	16.8	73.0	693	36	AA660414	AA660414 00290 MtR
c 2	15.8	68.7	539	22	H11322	H11322 ym13a03.s1
3	15.8	68.7	433	22	H13911	H13911 yj08g05.r1
c 4	15.8	68.7	281	23	H56483	H56483 yt87e12.s1
c 5	15.8	68.7	463	24	H98744	H98744 yx12c12.s1
6	15.8	68.7	421	25	W08531	W08531 mb47b04.r1
7	15.8	68.7	528	26	W41942	W41942 mc64e07.r1
c 8	15.8	68.7	447	26	W60228	W60228 zd28g09.s1
c 9	15.8	68.7	470	26	W76544	W76544 zd60d09.s1
10	15.8	68.7	324	27	AA003148	AA003148 mg51g01.r
11	15.8	68.7	432	27	AA010775	AA010775 ze22f04.r
12	15.8	68.7	497	27	AA036822	AA036822 zk29b05.r
13	15.8	68.7	360	27	AA038824	AA038824 mi95g07.r
c 14	15.8	68.7	603	27	AA057660	AA057660 zl93d09.s
15	15.8	68.7	430	27	W82322	W82322 mf04g11.r1
16	15.8	68.7	420	27	W82575	W82575 mf04a11.r1
17	15.8	68.7	390	28	AA059678	AA059678 mj75d09.r
18	15.8	68.7	495	28	AA060109	AA060109 mj71g12.r
19	15.8	68.7	542	28	AA061890	AA061890 mj92h08.r
c 20	15.8	68.7	490	28	AA075167	AA075167 zm86g04.s
21	15.8	68.7	553	28	AA082814	AA082814 zn25b04.r
22	15.8	68.7	417	28	AA103480	AA103480 mo24h04.r
c 23	15.8	68.7	454	29	AA132788	AA132788 zo22d08.s
24	15.8	68.7	415	29	AA145601	AA145601 mr63f03.r
c 25	15.8	68.7	586	29	AA147813	AA147813 zo48b09.s
26	15.8	68.7	437	29	AA151511	AA151511 zl26h07.r
27	15.8	68.7	336	29	AA183476	AA183476 mo97f03.r
c 28	15.8	68.7	555	29	AA187994	AA187994 zp67h06.s
c 29	15.8	68.7	442	30	AA206634	AA206634 zq80c11.s
30	15.8	68.7	563	30	AA222589	AA222589 mv70b12.r
c 31	15.8	68.7	476	30	AA243010	AA243010 zr25h03.s
32	15.8	68.7	485	30	AA250948	AA250948 zs06h11.r
33	15.8	68.7	334	30	AA269452	AA269452 va90a06.r
34	15.8	68.7	406	32	AA336931	AA336931 EST41591
c 35	15.8	68.7	521	33	AA447651	AA447651 zw97g04.s
36	15.8	68.7	544	34	AA472282	AA472282 vh01b03.r
37	15.8	68.7	372	34	AA474481	AA474481 vd55e09.r
38	15.8	68.7	367	34	AA475447	AA475447 vh15a11.r
39	15.8	68.7	545	34	AA516883	AA516883 vh88h03.r
c 40	15.8	68.7	552	34	AA521284	AA521284 aa75h07.s
41	15.8	68.7	423	35	AA555952	AA555952 vi64g01.r
c 42	15.8	68.7	584	35	AA594843	AA594843 no21b08.s
c 43	15.8	68.7	694	36	AA640937	AA640937 nr72e09.s
c 44	15.8	68.7	435	36	AA662313	AA662313 nu97d03.s
c 45	15.8	68.7	503	47	AI540076	AI540076 td09a05.x

ALIGNMENTS

RESULT 1
AA660414/c
LOCUS AA660414 693 bp mRNA EST 10-NOV-1997
DEFINITION 00290 MtRHE Medicago truncatula cDNA 5', mRNA sequence.
ACCESSION AA660414
NID g2604458
VERSION AA660414.1 GI:2604458
KEYWORDS EST.

```
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
            eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
            Medicago.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   Covitz,P.A., Smith,L.S. and Long,S.R.
  TITLE     Expressed sequence tags from a root hair-enriched Medicago
            truncatula cDNA library
  JOURNAL   Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1290520.

            Contact: Long SR
            Department of Biological Sciences and Howard Hughes Medical
            Institute
            Stanford University
            Gilbert Biology, Stanford, CA 94305-5020, USA
            Tel: 650 723 3232
            Fax: 650 725 8309
            Email: fa.srl@forsythe.stanford.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .693
                     /organism="Medicago truncatula"
                     /cultivar="Jemalong"
                     /db_xref="taxon:3880"
                     /clone_lib="MtRHE"
                     /tissue_type="Root hairs & tips"
                     /dev_stage="2-3 day old seedlings"
                     /note="Organ: Root; Vector: pBK-CMV; Site_1: EcoRI;
                     Site_2: XhoI; cDNA was synthesized from a pooled mRNA prep
                     from elongating root  hairs (30% w/w) and 2-3cm root tips
                     (70% w/w).  XhoI-oligo-dT  linker-primer and EcoRI
                     adaptors were used. cDNAs was cloned  unidirectionally
                     into lambda ZAP Express (Stratagene), amplified,  and
                     mass-excised into pBK-CMV vector plasmids.   More
                     information is available at http://bio-SRL8.stanford.edu."
BASE COUNT      187 a    149 c    133 g    178 t     46 others
ORIGIN


  Query Match             73.0%;  Score 16.8;  DB 36;  Length 693;
  Best Local Similarity   65.2%;  Pred. No. 38;
  Matches   15;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 gcngtngartggaayttygcnmg 23
          || || ||:| |||:||:|| :|
Db    206 GCAGTGGAGTTGAATTTTGCACG 184


RESULT   2
H11322/c
LOCUS       H11322         539 bp    mRNA            EST       26-JUN-1995
DEFINITION  ym13a03.s1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:47520 3', mRNA sequence.
ACCESSION   H11322
NID         g876142
VERSION     H11322.1  GI:876142
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     On May 10, 1995 this sequence version replaced gi:805432.

            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1450
            High quality sequence stops: 410
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1450   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 410.
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /db_xref="GDB:420061"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:47520"
                     /clone_lib="Soares infant brain 1NIB"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                     I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Hind III adaptors
                     (Pharmacia), digested with Not I and directionally cloned
                     into the Not I and Hind III sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      135 a    132 c    131 g    135 t      6 others
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 22;  Length 539;
  Best Local Similarity   63.6%;  Pred. No. 1.3e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnmg 23
          | || ||: ||||:||:|| :|
Db    453 CAGTTGAGGGGAACTTTGCCAG 432


RESULT   3
H13911
LOCUS       H13911         433 bp    mRNA            EST       27-JUN-1995
DEFINITION  yj08g05.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:148184 5', mRNA sequence.
ACCESSION   H13911
NID         g878731
VERSION     H13911.1  GI:878731
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 433)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT

            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
```

```
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 591
            High quality sequence stops: 331
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 591    Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 331.
FEATURES              Location/Qualifiers
     source           1. .433
                      /organism="Homo sapiens"
                      /db_xref="GDB:559920"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:148184"
                      /clone_lib="Soares placenta Nb2HP"
                      /sex="Female"
                      /dev_stage="placenta obtained at birth (full term)"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo.  "
BASE COUNT      103 a    105 c    120 g    103 t      2 others
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 22;  Length 433;
  Best Local Similarity   63.6%;  Pred. No. 1.2e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnmg 23
          | || ||: ||||:||:|| :|
Db    108 CAGTTGAGGGGAACTTTGCCAG 129


RESULT   4
H56483/c
LOCUS       H56483         281 bp    mRNA            EST       02-OCT-1995
DEFINITION  yt87e12.s1 Soares_pineal_gland_N3HPG Homo sapiens cDNA clone
            IMAGE:231310 3' similar to gb:X71973 PHOSPHOLIPID HYDROPEROXIDE
            GLUTHATIONE PEROXIDASE (HUMAN);, mRNA sequence.
ACCESSION   H56483
NID         g1005127
VERSION     H56483.1  GI:1005127
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 281)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     On May 9, 1995 this sequence version replaced gi:803053.

            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1524
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: polyT not found
            Insert Length: 1524    Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .281
                      /organism="Homo sapiens"
                      /db_xref="GDB:3861301"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:231310"
                      /clone_lib="Soares_pineal_gland_N3HPG"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: pineal gland; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                      1st strand cDNA was primed with a Not I - oligo(dT) primer
                      [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCGTTTTTTTTTTTTTTTTT
                      3'], double-stranded cDNA was size selected, ligated to
                      Eco RI adapters (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of a modified pT7T3
                      vector (Pharmacia). Library constructed by Bento Soares
                      and M.Fatima Bonaldo. "
BASE COUNT       45 a     81 c     92 g     53 t     10 others
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 23;  Length 281;
  Best Local Similarity   63.6%;  Pred. No. 1.2e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 gcngtngartggaayttygcnm 22
          ||  | ||:|||||:||:|| :
Db    271 GCCATAGAGTGGAACTTTGCCA 250


RESULT   5
H98744/c
LOCUS       H98744         463 bp    mRNA            EST       15-DEC-1995
DEFINITION  yx12c12.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
            IMAGE:261526 3', mRNA sequence.
ACCESSION   H98744
NID         g1123412
VERSION     H98744.1  GI:1123412
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 463)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     On Nov 22, 1995 this sequence version replaced gi:1071096.

            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 377
```

```
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1012   Std Error: 0.00
            Seq primer: m13 -40 forward
            High quality sequence stop: 377.
FEATURES              Location/Qualifiers
     source           1. .463
                      /organism="Homo sapiens"
                      /db_xref="GDB:3871168"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:261526"
                      /clone_lib="Soares melanocyte 2NbHM"
                      /sex="Male"
                      /tissue_type="melanocyte"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library constructed by Bento Soares and
                      M.Fatima Bonaldo. RNA from normal foreskin melanocytes
                      (FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT      123 a    122 c    110 g    107 t      1 others
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 24;  Length 463;
  Best Local Similarity   63.6%;  Pred. No. 1.2e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnmg 23
          | || ||: ||||:||:|| :|
Db    433 CAGTTGAGGGGAACTTTGCCAG 412


RESULT   6
W08531
LOCUS       W08531          421 bp    mRNA            EST       05-SEP-1996
DEFINITION  mb47b04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
            IMAGE:332527 5', mRNA sequence.
ACCESSION   W08531
NID         g1282533
VERSION     W08531.1  GI:1282533
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Nov 29, 1993 this sequence version replaced gi:430122.

            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:213927
            Seq primer: ETPrimer
            High quality sequence stop: 403.
FEATURES              Location/Qualifiers
     source           1. .421
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /map="1"
                      /clone="IMAGE:332527"
                      /clone_lib="Soares mouse p3NMF19.5"
                      /dev_stage="19.5 dpc total fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA was kindly provided by
                      Dr. Minoru Ko (Wayne State University)."
BASE COUNT      93 a    115 c    123 g     90 t
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 25;  Length 421;
  Best Local Similarity   63.6%;  Pred. No. 1.2e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnmg 23
          | || ||: ||||:||:|| :|
Db    104 CAGTTGAGGGGAACTTTGCCAG 125


RESULT   7
W41942
LOCUS       W41942          528 bp    mRNA            EST       20-MAY-1996
DEFINITION  mc64e07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:353316 5', mRNA sequence.
ACCESSION   W41942
NID         g1325656
VERSION     W41942.1  GI:1325656
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Apr 14, 1993 this sequence version replaced gi:693673.

            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:225116
            Seq primer: ETPrimer
            High quality sequence stop: 301.
FEATURES              Location/Qualifiers
     source           1. .528
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
```

```
                     /db_xref="taxon:10090"
                     /clone="IMAGE:353316"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo.  "
BASE COUNT      129 a    134 c    148 g    117 t
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 26;  Length 528;
  Best Local Similarity   63.6%;  Pred. No. 1.3e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnmg 23
          | || ||: ||||:||:|| :|
Db    141 CAGTTGAGGGGAACTTTGCCAG 162


RESULT   8
W60228/c
LOCUS       W60228         447 bp    mRNA            EST       15-OCT-1996
DEFINITION  zd28g09.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:342016 3', mRNA sequence.
ACCESSION   W60228
NID         g1366989
VERSION     W60228.1  GI:1366989
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     On Oct 18, 1995 this sequence version replaced gi:1023563.

            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 857   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 405.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Homo sapiens"
                     /db_xref="GDB:1267391"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:342016"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
BASE COUNT      117 a    116 c    104 g    108 t      2 others
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 26;  Length 447;
  Best Local Similarity   68.2%;  Pred. No. 1.2e+02;
  Matches   15;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnmg 23
          | |||||: ||||:||:|| :|
Db    443 CAGTNGAGGGGAACTTTGCCAG 422


RESULT   9
W76544/c
LOCUS       W76544         470 bp    mRNA            EST       16-OCT-1996
DEFINITION  zd60d09.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:345041 3', mRNA sequence.
ACCESSION   W76544
NID         g1386779
VERSION     W76544.1  GI:1386779
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     On Apr 14, 1993 this sequence version replaced gi:716583.

            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1070   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 414.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Homo sapiens"
                     /db_xref="GDB:1270416"
                     /db_xref="taxon:9606"
                     /map="5 q13; 754A01; 12; 12q12-12q13.13"
                     /clone="IMAGE:345041"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
```

/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

BASE COUNT 122 a 126 c 113 g 109 t
ORIGIN

Query Match 68.7%; Score 15.8; DB 26; Length 470;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 cngtngartggaayttygcnmg 23
| || ||: |||||:||:|| :|
Db 435 CAGTTGAGGGGAACTTTGCCAG 414

RESULT 10
AA003148
LOCUS AA003148 324 bp mRNA EST 19-JUL-1996
DEFINITION mg51g01.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:427344 5', mRNA sequence.
ACCESSION AA003148
NID g1446605
VERSION AA003148.1 GI:1446605
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 324)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Nov 4, 1993 this sequence version replaced.

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:261896
Seq primer: ETPrimer
High quality sequence stop: 287.

FEATURES Location/Qualifiers
source 1. .324
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:427344"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT 79 a 90 c 94 g 61 t
ORIGIN

Query Match 68.7%; Score 15.8; DB 27; Length 324;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 cngtngartggaayttygcnmg 23
| || ||: |||||:||:|| :|
Db 109 CAGTTGAGGGGAACTTTGCCAG 130

RESULT 11
AA010775
LOCUS AA010775 432 bp mRNA EST 29-NOV-1996
DEFINITION ze22f04.r1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:359743 5', mRNA sequence.
ACCESSION AA010775
NID g1471802
VERSION AA010775.1 GI:1471802
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 432)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On Apr 14, 1993 this sequence version replaced gi:785787.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 697 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 346.

FEATURES Location/Qualifiers
source 1. .432
/organism="Homo sapiens"
/db_xref="GDB:1276287"
/db_xref="taxon:9606"
/map="4 q35-qter"
/clone="IMAGE:359743"
/clone_lib="Soares_fetal_heart_NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."
BASE COUNT 108 a 113 c 105 g 106 t
ORIGIN

Query Match 68.7%; Score 15.8; DB 27; Length 432;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 cngtngartggaayttygcnmg 23
| || ||: ||||:||:|| :|
Db 215 CAGTTGAGGGGAACTTTGCCAG 236

RESULT 12
AA036822
LOCUS AA036822 497 bp mRNA EST 26-AUG-1996
DEFINITION zk29b05.r1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone IMAGE:471921 5', mRNA sequence.
ACCESSION AA036822
NID g1509860
VERSION AA036822.1 GI:1509860
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 497)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On Apr 14, 1993 this sequence version replaced gi:785518.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 425.
FEATURES Location/Qualifiers
source 1. .497
/organism="Homo sapiens"
/db_xref="GDB:3757923"
/db_xref="taxon:9606"
/clone="IMAGE:471921"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."
BASE COUNT 123 a 125 c 131 g 114 t 4 others
ORIGIN

Query Match 68.7%; Score 15.8; DB 27; Length 497;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 cngtngartggaayttygcnmg 23
| || ||: ||||:||:|| :|
Db 172 CAGTTGAGGGGAACTTTGCCAG 193

RESULT 13
AA038824
LOCUS AA038824 360 bp mRNA EST 28-AUG-1996
DEFINITION mi95g07.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone IMAGE:474396 5' similar to WP:C28H8.12 CE01823 ;, mRNA sequence.
ACCESSION AA038824
NID g1514249
VERSION AA038824.1 GI:1514249
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 360)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 21, 1992 this sequence version replaced gi:276467.

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:285140
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 353.
FEATURES Location/Qualifiers
source 1. .360
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="8"
/clone="IMAGE:474396"
/clone_lib="Soares mouse p3NMF19.5"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."
BASE COUNT 95 a 97 c 97 g 71 t
ORIGIN

Query Match 68.7%; Score 15.8; DB 27; Length 360;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;

```
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnmg 23
          | || ||: |||||:||:|| :|
Db    299 CAGTTGAGGGGAACTTTGCCAG 320


RESULT  14
AA057660/c
LOCUS       AA057660       603 bp    mRNA            EST       19-MAY-1997
DEFINITION  zl93d09.s1 Stratagene corneal stroma (#937222) Homo sapiens cDNA
            clone IMAGE:512177 3' similar to TR:G1255188 G1255188 DYNAMITIN. ;,
            mRNA sequence.
ACCESSION   AA057660
NID         g1550300
VERSION     AA057660.1  GI:1550300
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 603)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1297673.

            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1942   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1. .603
                     /organism="Homo sapiens"
                     /db_xref="GDB:3844673"
                     /db_xref="taxon:9606"
                     /map="12"
                     /clone="IMAGE:512177"
                     /clone_lib="Stratagene corneal stroma (#937222)"
                     /dev_stage="76 years"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: cornea; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Corneal fibroblasts grown from explants, 76
                     years. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      150 a    150 c    153 g    148 t      2 others
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 27;  Length 603;
  Best Local Similarity   63.6%;  Pred. No. 1.3e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnmg 23
          | || ||: |||||:||:|| :|
Db    439 CAGTTGAGGGGAACTTTGCCAG 418


RESULT  15
W82322
LOCUS       W82322         430 bp    mRNA            EST       12-SEP-1996
DEFINITION  mf04g11.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
            IMAGE:404132 5' similar to WP:C28H8.12 CE01823 ;, mRNA sequence.
ACCESSION   W82322
NID         g1540001
VERSION     W82322.1  GI:1540001
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1393548.

            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:247900
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 393.
FEATURES             Location/Qualifiers
     source          1. .430
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:404132"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
BASE COUNT      113 a    112 c    118 g     87 t
ORIGIN


  Query Match             68.7%;  Score 15.8;  DB 27;  Length 430;
  Best Local Similarity   63.6%;  Pred. No. 1.2e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 cngtngartggaayttygcnmg 23
          | || ||: |||||:||:|| :|
Db    371 CAGTTGAGGGGAACTTTGCCAG 392


Search completed: October  1, 1999, 15:03:38
Job time: 4382 sec
```